

Linear Algebra and Applications: An Inquiry-Based Approach

Feryal Alayont
Steven Schlicker

Grand Valley State University



Contents

Preface	xvii
A Free and Open-Source Linear Algebra Text	xvii
Goals	xvii
Layout	xviii
Acknowledgements	xix
To the Student	xx
I Systems of Linear Equations	1
1 Introduction to Systems of Linear Equations	3
Application: Electrical Circuits	3
Introduction	4
Notation and Terminology	6
Solving Systems of Linear Equations	7
The Geometry of Solution Sets of Linear Systems	8
Examples	10
Summary	13
Exercises	13
True/False Questions	16
Project: Modeling an Electrical Circuit and the Wheatstone Bridge Circuit	17
2 The Matrix Representation of a Linear System	21
Application: Simpson's Rule	21
Introduction	22
Simplifying Linear Systems Represented in Matrix Form	24
Linear Systems with Infinitely Many Solutions	26

Linear Systems with No Solutions	28
Examples	29
Summary	33
Exercises	33
Project: A Polynomial Fitting Application: Simpson's Rule	35
3 Row Echelon Forms	39
Application: Balancing Chemical Reactions	39
Introduction	40
The Echelon Forms of a Matrix	41
Determining the Number of Solutions of a Linear System	44
Producing the Echelon Forms	45
Examples	48
Summary	51
Exercises	53
Project: Modeling a Chemical Reaction	55
4 Vector Representation	57
Application: The Knight's Tour	57
Introduction	58
Vectors and Vector Operations	61
Geometric Representation of Vectors and Vector Operations	63
Linear Combinations of Vectors	64
The Span of a Set of Vectors	67
Examples	69
Summary	72
Exercises	73
Project: Analyzing Knight Moves	75
5 The Matrix-Vector Form of a Linear System	79
Application: Modeling an Economy	79
Introduction	80
The Matrix-Vector Product	83
The Matrix-Vector Form of a Linear System	84
Homogeneous and Nonhomogeneous Systems	87

The Geometry of Solutions to the Homogeneous System	89
Examples	90
Summary	93
Exercises	94
Project: Input-Output Models	97
6 Linear Dependence and Independence	103
Application: Bézier Curves	103
Introduction	104
Linear Independence	107
Determining Linear Independence	111
Minimal Spanning Sets	111
Examples	116
Summary	118
Exercises	119
Project: Generating Bézier Curves	121
7 Matrix Transformations	125
Application: Computer Graphics	125
Introduction	126
Properties of Matrix Transformations	128
Onto and One-to-One Transformations	129
Examples	132
Summary	134
Exercises	135
Project: The Geometry of Matrix Transformations	138
II Matrices	141
8 Matrix Operations	143
Application: Algorithms for Matrix Multiplication	143
Introduction	144
Properties of Matrix Addition and Multiplication by Scalars	146
A Matrix Product	147
The Transpose of a Matrix	153



Properties of the Matrix Transpose	154
Examples	155
Summary	160
Exercises	161
Project: Strassen's Algorithm and Partitioned Matrices	163
9 Introduction to Eigenvalues and Eigenvectors	169
Application: The Google PageRank Algorithm	169
Introduction	170
Eigenvalues and Eigenvectors	171
Dynamical Systems	173
Examples	175
Summary	180
Exercises	180
Project: Understanding the PageRank Algorithm	182
10 The Inverse of a Matrix	189
Application: Modeling an Arms Race	189
Introduction	190
Invertible Matrices	191
Finding the Inverse of a Matrix	193
Properties of the Matrix Inverse	194
Examples	195
Summary	197
Exercises	198
Project: The Richardson Arms Race Model	199
11 The Invertible Matrix Theorem	203
Introduction	203
The Invertible Matrix Theorem	204
Examples	208
Summary	209
Exercises	209



III The Vector Space \mathbb{R}^n	211
12 The Structure of \mathbb{R}^n	213
Application: Connecting GDP and Consumption in Romania	213
Introduction	215
Vector Spaces	216
The Subspace Spanned by a Set of Vectors	219
Examples	221
Summary	223
Exercises	224
Project: Least Squares Linear Approximation	227
13 The Null Space and Column Space of a Matrix	231
Application: The Lights Out Game	231
Introduction	232
The Null Space of a Matrix and the Kernel of a Matrix Transformation	233
The Column Space of a Matrix and the Range of a Matrix Transformation	234
The Row Space of a Matrix	235
Bases for $\text{Nul } A$ and $\text{Col } A$	236
Examples	238
Summary	241
Exercises	242
Project: Solving the Lights Out Game	243
14 Eigenspaces of a Matrix	249
Application: Population Dynamics	249
Introduction	250
Eigenspaces of Matrix	251
Linearly Independent Eigenvectors	252
Examples	253
Summary	257
Exercises	257
Project: Modeling Population Migration	258
15 Bases and Dimension	261
Application: Lattice Based Cryptography	261



Introduction	262
The Dimension of a Subspace of \mathbb{R}^n	263
Conditions for a Basis of a Subspace of \mathbb{R}^n	265
Finding a Basis for a Subspace	266
Rank of a Matrix	266
Examples	267
Summary	269
Exercises	270
Project: The GGH Cryptosystem	272
IV Eigenvalues and Eigenvectors	277
16 The Determinant	279
Application: Area and Volume	279
Introduction	280
The Determinant of a Square Matrix	281
Cofactors	282
The Determinant of a 3×3 Matrix	285
Two Devices for Remembering Determinants	286
Examples	286
Summary	290
Exercises	290
Project: Area and Volume Using Determinants	291
17 The Characteristic Equation	297
Application: Modeling the Second Law of Thermodynamics	297
Introduction	298
The Characteristic Equation	299
Eigenspaces, A Geometric Example	300
Dimensions of Eigenspaces	301
Examples	302
Summary	305
Exercises	306
Project: The Ehrenfest Model	307

18 Diagonalization	313
Application: The Fibonacci Numbers	313
Introduction	314
Diagonalization	315
Similar Matrices	316
Similarity and Matrix Transformations	317
Diagonalization in General	318
Examples	320
Summary	324
Exercises	324
Project: Binet's Formula for the Fibonacci Numbers	327
19 Approximating Eigenvalues and Eigenvectors	331
Application: Leslie Matrices and Population Modeling	331
Introduction	332
The Power Method	332
The Inverse Power Method	336
Examples	337
Summary	339
Exercises	340
Project: Managing a Sheep Herd	342
20 Complex Eigenvalues	349
Application: The Gershgorin Disk Theorem	349
Introduction	350
Complex Eigenvalues	350
Rotation and Scaling Matrices	351
Matrices with Complex Eigenvalues	352
Examples	354
Summary	355
Exercises	356
Project: Understanding the Gershgorin Disk Theorem	357
21 Properties of Determinants	363
Introduction	363



Elementary Row Operations and Their Effects on the Determinant	365
Elementary Matrices	365
Geometric Interpretation of the Determinant	368
An Explicit Formula for the Inverse and Cramer's Rule	369
The Determinant of the Transpose	372
Row Swaps and Determinants	373
Cofactor Expansions	375
The LU Factorization of a Matrix	376
Examples	379
Summary	382
Exercises	382
V Vector Spaces	385
22 Vector Spaces	387
Application: The Hat Puzzle	387
Introduction	388
Spaces with Similar Structure to \mathbb{R}^n	390
Vector Spaces	391
Subspaces	394
Examples	396
Summary	399
Exercises	400
Project: Hamming Codes and the Hat Puzzle	403
23 Bases for Vector Spaces	407
Application: Image Compression	407
Introduction	409
Linear Independence	411
Bases	412
Finding a Basis for a Vector Space	413
Examples	414
Summary	416
Exercises	417
Project: Image Compression with Wavelets	419

24 The Dimension of a Vector Space	427
Application: Principal Component Analysis	427
Introduction	428
Finite Dimensional Vector Spaces	428
The Dimension of a Subspace	429
Conditions for a Basis of a Vector Space	429
Examples	430
Summary	433
Exercises	433
Project: Understanding Principal Component Analysis	435
25 Coordinate Vectors and Coordinate Transformations	443
Application: Calculating Sums	443
Introduction	444
Bases as Coordinate Systems	445
The Coordinate Transformation	447
Examples	450
Summary	452
Exercises	452
Project: Finding Formulas for Sums of Powers	455
26 Change of Basis	459
Application: Describing Orbits of Planets	459
Introduction	459
The Change of Basis Matrix	461
Finding a Change of Basis Matrix	462
Properties of the Change of Basis Matrix	464
Examples	465
Summary	467
Exercises	468
Project: Planetary Orbits and Change of Basis	470
VI Inner Product Spaces	475
27 The Dot Product in \mathbb{R}^n	477



Application: Hidden Figures in Computer Graphics	477
Introduction	478
The Distance Between Vectors	481
The Angle Between Two Vectors	481
Orthogonal Projections	483
Orthogonal Complements	484
Examples	486
Summary	488
Exercises	489
Project: Back-Face Culling	491
28 Orthogonal and Orthonormal Bases in \mathbb{R}^n	495
Application: Rotations in 3D	495
Introduction	496
Orthogonal Sets	496
Properties of Orthogonal Bases	498
Orthonormal Bases	498
Orthogonal Matrices	499
Examples	501
Summary	502
Exercises	503
Project: Understanding Rotations in 3-Space	504
29 Inner Products	509
Application: Fourier Series	509
Introduction	510
Inner Product Spaces	511
The Length of a Vector	512
Orthogonality in Inner Product Spaces	513
Orthogonal and Orthonormal Bases in Inner Product Spaces	513
Orthogonal Projections onto Subspaces	515
Best Approximations	517
Orthogonal Complements	519
Examples	520
Summary	521



Exercises	523
Project: Fourier Series and Musical Tones	526
30 The Gram-Schmidt Process	531
Application: Gaussian Quadrature	531
Introduction	531
The Gram-Schmidt Process	532
The QR Factorization of a Matrix	534
Examples	536
Summary	539
Exercises	540
Project: Gaussian Quadrature and Legendre Polynomials	542
VII Orthogonality	547
31 Orthogonal Diagonalization	549
Application: The Multivariable Second Derivative Test	549
Introduction	550
Symmetric Matrices	551
The Spectral Decomposition of a Symmetric Matrix A	555
Examples	557
Summary	559
Exercises	559
Project: The Second Derivative Test for Functions of Two Variables	561
32 Quadratic Forms and the Principal Axis Theorem	565
Application: The Tennis Racket Effect	565
Introduction	566
Equations Involving Quadratic Forms in \mathbb{R}^2	567
Classifying Quadratic Forms	569
Examples	571
Summary	574
Exercises	574
Project: The Tennis Racket Theorem	577



33 The Singular Value Decomposition	581
Application: Search Engines and Semantics	581
Introduction	582
The Operator Norm of a Matrix	582
The SVD	586
SVD and the Null, Column, and Row Spaces of a Matrix	590
Examples	590
Summary	594
Exercises	595
Project: Latent Semantic Indexing	597
34 Approximations Using the Singular Value Decomposition	603
Application: Global Positioning System	603
Introduction	604
Image Compression	604
Calculating the Error in Approximating an Image	607
The Condition Number of a Matrix	609
Pseudoinverses	612
Least Squares Approximations	616
Examples	620
Summary	625
Exercises	626
Project: GPS and Least Squares	630
VIII Linear Transformations	635
35 Linear Transformations	637
Application: Fractals	637
Introduction	638
Onto and One-to-One Transformations	639
The Kernel and Range of Linear Transformation	640
Isomorphisms	641
Examples	642
Summary	644
Exercises	645

Project: Fractals via Iterated Function Systems	648
36 The Matrix of a Linear Transformation	655
Application: Secret Sharing Algorithms	655
Introduction	656
Linear Transformations from \mathbb{R}^n to \mathbb{R}^m	656
The Matrix of a Linear Transformation	657
A Connection between $\text{Ker}(T)$ and a Matrix Representation of T	659
Examples	659
Summary	662
Exercises	663
Project: Shamir's Secret Sharing and Lagrange Polynomials	666
37 Eigenvalues of Linear Transformations	671
Application: Linear Differential Equations	671
Introduction	672
Finding Eigenvalues and Eigenvectors of Linear Transformations	672
Diagonalization	673
Examples	674
Summary	677
Exercises	677
Project: Linear Transformations and Differential Equations	678
A Complex Numbers	683
Complex Numbers	683
Conjugates and Modulus	685
B Answers and Hints to Selected Exercises	687

Preface

A Free and Open-Source Linear Algebra Text

Mathematics is for everyone – whether as a gateway to other fields or as background for higher level mathematics. With linear algebra gaining importance in many applications, we feel that access to or the cost of a textbook should not stand in the way of a successful experience in learning linear algebra. Therefore, we made our textbook available to everyone for free download for their own non-commercial use. We especially encourage its use in linear algebra classrooms for instructors who are looking for an inquiry-based textbook or a supplemental resource to accompany their course. If an instructor would like to make changes to any of the files to better suit your students' needs, we offer source files for the text by making a request to the authors.

This work is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License. The graphic



that appears throughout the text shows that the work is licensed with the Creative Commons, that the work may be used for free by any party so long as attribution is given to the author(s), that the work and its derivatives are used in the spirit of “share and share alike,” and that no party may sell this work or any of its derivatives for profit. Full details may be found by visiting

<http://creativecommons.org/licenses/by-nc-sa/3.0/>

or sending a letter to Creative Commons, 444 Castro Street, Suite 900, Mountain View, California, 94041, USA.

Goals

Linear Algebra and Applications: An Inquiry-Based Approach provides a novel inquiry-based learning approach to linear algebra, as well as incorporating aspects of an inverted classroom. The impetus for this book lies in our approach to teaching linear algebra. We place an emphasis on active learning and on developing students' intuition through their investigation of examples. For us, active learning involves students – they are DOING something instead of being passive learners.

What students are doing when they are actively learning might include discovering, processing, discussing, applying information, writing intensive assignments, engaging in common intellectual in-class experiences or collaborative assignments and projects. Although it is difficult to capture the essence of active learning in a textbook, this book is our attempt to do just that.

Our goals for these materials are several.

- To carefully introduce the ideas behind the definitions and theorems to help students develop intuition and understand the logic behind them.
- To help students understand that mathematics is not done as it is often presented. We expect students to experiment through examples, make conjectures, and then refine their conjectures. We believe it is important for students to learn that definitions and theorems don't pop up completely formed in the minds of most mathematicians, but are the result of much thought and work.
- To help students develop their communication skills in mathematics. We expect our students to read and complete activities before class and come prepared with questions. While in class, students work to discover many concepts on their own through guided activities. Of course, students also individually write solutions to exercises on which they receive significant feedback. Communication skills are essential in any discipline and we place a heavy focus on their development.
- To have students actively involved in each of these items through in-class and out-of-class activities, in-class presentations (this is of course up to the instructor), and problem sets.

Layout

This text is formatted into sections, each of which contains preview activities, in-class activities, worked examples, and exercises. Most sections conclude with an application project – an application of the material in the section. The various types of activities serve different purposes.

- Preview activities are designed for students to complete before class to motivate the upcoming topic and prepare them with the background and information they need for the class activities and discussion.
- We generally use the regular activities to engage students during class in critical thinking experiences. These activities are used to provide motivation for the material, opportunities for students to develop course material on their own, or examples to help reinforce the meanings of definitions or theorems. The ultimate goal is to help students develop their intuition for and understanding of linear algebra concepts.
- Worked examples are included in each section. Part of the philosophy of this text is that students will develop the tools and understanding they need to work homework assignments through the preview, in-class activities, and class discussions. But some students express a desire for fully-worked examples in the text to reference during homework. In order to preserve the flow of material within a section, we include worked examples at the end of each section.



- The Linear Algebra Curriculum Study Group (LACSG) was created in 1990 to “initiate substantial and sustained national interest in improving the undergraduate linear algebra curriculum”.¹ Funded by the national science foundation, the group formed a broad based panel of faculty, and in consultation with client disciplines they produced a document that makes recommendations for the linear algebra curriculum. One of the recommendations, “Mathematics departments should seriously consider making their first course in linear algebra a matrix-oriented course.” is followed in this text. They suggest that this approach “implies less emphasis on abstraction and more emphasis on problem solving and motivating applications.” and that “Some applications of linear algebra should be included to give an indication of the pervasive use of linear algebra in many client disciplines. Such applications necessarily will be limited by the need to minimize technical jargon and information from outside the course. Students should see the course as one of the most potentially useful mathematics courses they will take as an undergraduate.” Also, the 2015 Committee on the Undergraduate Program in Mathematics (CUPM) report to the Mathematical Association of America² recommends that “Every Linear Algebra course should incorporate interesting applications, both to highlight the broad usefulness of linear algebra and to help students see the role of the theory in the subject as it is applied. Attractive applications may also entice students majoring in other disciplines to choose a minor or additional major in mathematics.” All but two sections in this text include a substantial application. Each section begins with a short description of an application that uses the material from the section, then concludes with a project that develops the application in more detail. (The two sections that do not include projects are sections that are essentially long proofs – one section that contains formal proofs of the equivalences of the different parts of the Invertible Matrix Theorem, and the other contains algebraic proofs of the properties of the determinant.) The projects are independent of the material in the text – the text can be used without the applications. The applications are written in such a way that they could also be used with other textbooks. The projects are written following an inquiry-based style similar to the text, with important parts of the applications developed through activities. So the projects can be assigned outside of class as independent work for students. Several of the projects are accompanied by GeoGebra applets or Sage worksheets which are designed to help the students better understand the applications.
- Each investigation contains a collection of exercises. The exercises occur at a variety of levels of difficulty and most force students to extend their knowledge in different ways. While there are some standard, classic problems that are included in the exercises, many problems are open ended and expect a student to develop and then verify conjectures.

Acknowledgements

The many beautiful images that are in this text were created by our colleague David Austin, to whom we owe a debt of gratitude. Our institution, Grand Valley State University, supported our

¹Their report is found in “The Linear Algebra Curriculum Study Group Recommendations for the First Course in Linear Algebra”, David Carlson, Charles R. Johnson, David C. Lay and A. Duane Porter, *The College Mathematics Journal*, Vol. 24, No. 1 (Jan., 1993), pp. 41-46.

²2015 CUPM Curriculum Guide to Majors in the Mathematical Sciences, Carol S. Schumacher and Martha J. Siegel, Co-Chairs, Paul Zorn, Editor. <https://www.maa.org/sites/default/files/CUPM%20Guide.pdf>



development of these materials early on through an internal university grant. Finally, we wish to thank our GVSU colleagues for their support and for being role models in the open source textbook movement.

To the Student

The inquiry-based format of this book means that you can be in charge of your own learning. The guidance of your instructor and support of your fellow classmates can help you develop your understanding of the topic. Your learning will benefit best if you engage in the material by completing all preview and in-class activities in order to fully develop your own intuition and understanding of the material, which can be achieved only by reflecting on the ideas and concepts and communicating your reflections with others. Don't be afraid to ask questions or to be wrong at times, for this is how people learn. Good luck! We will be happy to hear your comments about the book.





Part I

Systems of Linear Equations

Section 1

Introduction to Systems of Linear Equations

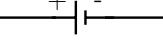
Focus Questions

By the end of this section, you should be able to give precise and thorough answers to the questions listed below. You may want to keep these questions in mind to focus your thoughts as you complete the section.

- What is a linear equation?
- What is a system of linear equations?
- What is a solution set of a system of linear equations?
- What are equivalent systems of linear equations?
- What operations can we use to solve a system of linear equations?

Application: Electrical Circuits

Linear algebra is concerned with the study of systems of linear equations. There are two important aspects to linear systems. One is to use given information to set up a system of equations that represents the information (this is called *modeling*), and the other is to solve the system. As an example of modeling, we consider the application to the very simple electrical circuit. An electrical circuit consists of

- one or more electrical sources, denoted by 
- one or more resistors, denoted by .

A source is a power supply like a battery, and a resistor is an object that consumes the electricity, like a lamp or a computer. A simple circuit consists of one or more sources connected to resistors,

like the one shown in Figure 1.2. The straight lines in the circuit indicate wires through which current flows. The points labeled P and Q are called *junctions* or *nodes*.

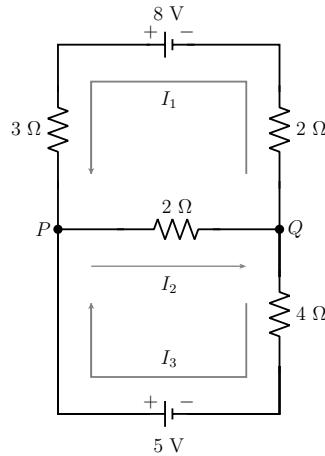


Figure 1.1: A circuit.

The source creates a charge that produces potential energy E measured in volts (V). Current flows out of the positive terminal of a source and runs through each branch of the circuit. Let I_1 , I_2 , and I_3 be the currents as illustrated in Figure 1.2. The goal is to find the current flowing in each branch of the circuit.

Linear algebra comes into play when analyzing a circuit based on the relationship between current I , resistance R , and voltage E . There are laws governing electrical circuits that state that $E = IR$ across a resistor. Additionally, Kirchoff's Current and Voltage Laws indicate how current behaves within the whole circuit. Using all these laws together, we derive the system

$$\begin{aligned} I_1 - I_2 + I_3 &= 0 \\ 5I_1 + 2I_2 &= 8 \\ 2I_2 + 4I_3 &= 5, \end{aligned}$$

where I_1 , I_2 , and I_3 are the currents at the points indicated in Figure 1.2. To finish analyzing the circuit, we now need to solve this system. In this section we will begin to learn systematic methods for solving systems of linear equations. More details about the derivation of these circuit equations can be found at the end of this section.

Introduction

Systems of linear equations arise in almost every field of study: mathematics, statistics, physics, chemistry, biology, economics, sociology, computer science, engineering, and many, many others. We will study the theory behind solving systems of linear equations, implications of this theory, and applications of linear algebra as we proceed throughout this text.

Preview Activity 1.1.



- (1) Consider the following system of two linear equations in two unknowns, x_1, x_2 :

$$\begin{aligned} 2x_1 - 3x_2 &= 0 \\ x_1 - x_2 &= 1. \end{aligned}$$

One way to solve such a system of linear equations is the method of substitution (where one equation is solved for one variable and then the resulting expression is substituted into the remaining equations). This method works well for simple systems of two equations in two unknowns, but becomes complicated if the number or complexity of the equations is increased.

Another method is elimination – the method that we will adopt in this book. Recall that the elimination method works by multiplying each equation by a suitable constant so that the coefficients of one of the variables in each equation is the same. Then we subtract corresponding sides of these equations to eliminate that variable.

Use the method of elimination to show that this system has the unique solution $x_1 = 3$ and $x_2 = 2$. Explain the specific steps you perform when using elimination.

- (2) Recall that a linear equation in two variables can be represented as a line in \mathbb{R}^2 , the Cartesian plane, where one variable corresponds to the horizontal axis and the other to the vertical axis. Represent the two equations $2x_1 - 3x_2 = 0$ and $x_1 - x_2 = 1$ in \mathbb{R}^2 and illustrate the solution to the system in your picture.
- (3) The previous example should be familiar to you as a system of two equations in two unknowns. Now we consider a system of three equations in three unknowns

$$I_1 - I_2 + I_3 = 0 \quad (1.1)$$

$$5I_1 + 2I_2 = 8 \quad (1.2)$$

$$2I_2 + 4I_3 = 5. \quad (1.3)$$

that arises from our electrical circuit in Figure 1.2, with currents I_1, I_2 , and I_3 as indicated in the circuit. In the remainder of this preview activity we will apply the method of elimination to solve the system of linear equations (1.1), (1.2), and (1.3).

- (a) Replace equation (1.2) with the new equation obtained by multiplying both sides of equation (1.1) by 5 and then subtracting corresponding sides of this equation from the appropriate sides of equation (1.2). Show that the resulting system is

$$\begin{aligned} I_1 - I_2 + I_3 &= 0 \\ 7I_2 - 5I_3 &= 8 \\ 2I_2 + 4I_3 &= 5. \end{aligned} \quad (1.4)$$

- (b) Now eliminate the variable I_2 from the last two equations in the system in part (a) by using equations (1.3) and (1.4) to show that $I_3 = 0.5$. Explain your process.
- (c) Once you know the value for I_3 , how can you find I_2 ? Then how do you find I_1 ? Use your method to show that the solution to this system is the ordered triple $(1, 1.5, 0.5)$. Interpret the result in terms of currents.



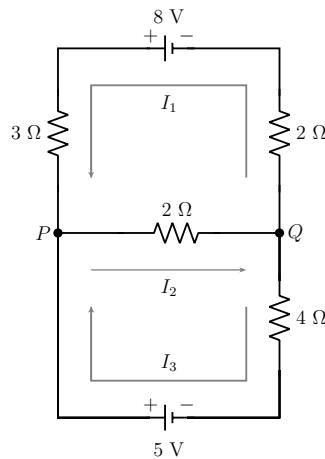


Figure 1.2: A circuit.

Notation and Terminology

To study linear algebra, we will need to agree on some general notation and terminology to represent our systems.

An equation like $4x_1 + x_2 = 8$ is called a linear equation because the variables (x_1 and x_2 in this case) are raised to the first power, and there are no products of variables. The equation $4x_1 + x_2 = 8$ is a linear equation in two variables, but we can make a linear equation with any number of variables we like.

Definition 1.1. A **linear equation** in the variables x_1, x_2, \dots, x_n is an equation of the form

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n = b,$$

where n is a positive integer and a_1, a_2, \dots, a_n and b are constants. The constants a_1, a_2, \dots, a_n are called the **coefficients** of the equation.

We can use any labels for the variables in a linear equation that we like, e.g., I_1, x_1, t_1 , and you should become comfortable working with variables in any form. We will usually use subscripts, as in x_1, x_2, x_3, \dots , to represent the variables as this notation allows us to have any number of variables. Other examples of linear equations are

$$x + 2y = 4 \quad \text{and} \quad \sqrt{2}x_1 - 3x_2 = \frac{1}{4}x_3 + \pi.$$

On the other hand, the equations

$$\frac{1}{x} + y - z = 0 \quad \text{and} \quad 2x_1 = \sqrt{x_2} - 5$$

are non-linear equations.

Definition 1.2. A **system of linear equations** is a collection of one or more linear equations in the same variables.

For example, the two equations

$$\begin{aligned}x_1 - x_2 &= 1 \\2x_1 + x_2 &= 5\end{aligned}\tag{1.5}$$

form a system of two linear equations in variables x_1, x_2 .

Definition 1.3. A **solution** to a system of linear equations is an ordered n -tuple (s_1, s_2, \dots, s_n) of numbers so that we obtain all true statements in the system when we replace the variable in order with s_1, s_2, \dots , and s_n .

For example, $x_1 = 2, x_2 = 1$, or simply $(2, 1)$, is a solution to the above system of linear equations in (1.5) as can be checked by substituting the variables into each equation. In solving a system of linear equations, we are interested in finding the set of all solutions, which we will call the *solution set of the system*. For the above system in (1.5), the solution set is the set containing the single point $(2, 1)$, denoted $\{(2, 1)\}$, because there is only one solution. If we consider just the equation $x_1 - x_2 = 0$ as our system, the solution set is the line $x_1 = x_2$ in the plane. More generally, a set of solutions is a collection of ordered n -tuples of numbers. We denote the set of all ordered n -tuples of numbers as \mathbb{R}^n . So, for example, \mathbb{R}^2 is the set of all ordered pairs, or just the standard coordinate plane, and \mathbb{R}^3 is the set of all ordered triples, or the three-dimensional space.

Solving Systems of Linear Equations

In Preview Activity 1.1, we were introduced to linear systems and the method of elimination for a system of two or three variables. Our goal now is to come up with a systematic method that will reduce any linear system to one that is easy to solve without changing the solution set of the system. Two linear systems will be called *equivalent* if they have the same solution set.

The operations we used in Preview Activity 1.1 to systematically eliminate variables so that we can solve a linear system are called *elementary operations on a system of linear equations* or just *elementary operations*. In the exercises you will argue that elementary operations do not change the solution set to a system of linear equations, a fact that is summarized in the following theorem.

Theorem 1.4. *The elementary operations on a system of linear equations:*

- (1) replacing one equation by the sum of that equation and a scalar multiple of another equation;
- (2) interchanging two equations;
- (3) replacing an equation by a nonzero scalar multiple of itself;

do not change the solution set to the system of equations.

When we apply these elementary operations our ultimate goal is to produce a system of linear equations in a simplified form with the same solution set, where the number of variables eliminated from the equations increase as we move from top to bottom. This method is called the *elimination* method.



Activity 1.1. For systems of linear equations with a small number of variables, many different methods could be used to find a solution. However, when a system gets large, ad-hoc methods become unwieldy. One of our goals is to develop an algorithmic approach to solving systems of linear equations that can be programmed and applied to any linear system, so we want to work in a very prescribed method as indicated in this activity. Ultimately, once we understand how the algorithm works, we will use calculators/computers to do the work. Apply the elimination method as described to show that the solution set of the following system is $(2, -1, 1)$.

$$\begin{aligned}x_1 + x_2 - x_3 &= 0 \\2x_1 + x_2 - x_3 &= 2 \\x_1 - x_2 + 2x_3 &= 5.\end{aligned}$$

- (a) Use the first equation to eliminate the variable x_1 in the second and third equations.
- (b) Use the new second equation to eliminate the variable x_2 in the third equation and find the value of x_3 .
- (c) Find values of x_2 and then x_1 .

Important Note: Technically, we don't really add two equations or multiply an equation by a scalar. When we refer to a scalar multiple of an equation, we mean the equation obtained by equating the scalar multiple of the expression on the left side of the equation and the same scalar multiple of the expression on the right side of the equation. Similarly, when we refer to a sum of two equations, we don't really add the equations themselves. Instead, we mean the equation obtained by equating the sum of the expressions on the left sides of the equations to the sum of the expressions on the right sides of the equations. We will use the terminology “scalar multiple of an equation” and “sum of two equations” as shorthand to mean what is described here.

Another Important Note: There is an important and subtle point to consider here. When we use these operations to find a solution to a system of equations, we are assuming that the system has a solution. The application of these operations then tells us what a solution must look like. However, there is no guarantee that the outcome is actually a solution – to be safe we should check to make sure that our result is a solution to the system. In the case of linear systems, though, every one of our operations on equations is reversible (if applied correctly), so the result will always be a solution (but this is not true in general for non-linear systems).

Terminology: A system of equations is called *consistent* if the system has at least one solution. If a system has no solutions, then it is said to be *inconsistent*.

The Geometry of Solution Sets of Linear Systems

We are familiar with linear equations in two variables from basic algebra and calculus (through linear approximations). The set of solutions to a system of linear equations in two variables has some geometry connected to it.



Activity 1.2. Recall that we examined the geometry of the system

$$\begin{aligned} 2x_1 - 3x_2 &= 0 \\ x_1 - x_2 &= 1 \end{aligned}$$

in Preview Activity 1.1 to show that the resulting solution set consists of a single point in the plane.

In this activity we examine the geometry of the system

$$\begin{aligned} 2x_1 - x_2 &= 1 \\ 2x_1 - 2x_2 &= 2. \end{aligned} \tag{1.6}$$

- (a) Consider the linear equation $2x_1 - 2x_2 = 2$ (or, equivalently $2x - 2y = 2$). What is the graph of the solution set (the set of points (x_1, x_2) satisfying this equation) of this single equation in the plane? Draw the graph to illustrate.
- (b) How can we represent the solution set of the system (1.6) of two equations graphically? How is this solution set related to the solution set of the single equation $2x_1 - 2x_2 = 2$? Why? How many solutions does the system (1.6) have?
- (c) There are exactly three possibilities for the number of solutions to a general system of two linear equations in two unknowns. Describe the geometric representations of solution sets for each of the possibilities. Illustrate each with a specific example (of your own) using a system of equations and sketching its geometric representation.

Activity 1.2 shows that there are three options for the solution set of a system: A system can have no solutions, one solution, or infinitely many solutions.

Now we consider systems of three variables. As an example, let us look at the linear equation $x + y + z = 1$ in the three variables x , y , and z . Notice that the points $(1, 0, 0)$, $(0, 1, 0)$, and $(0, 0, 1)$ all satisfy this equation. As a linear equation, the graph of $x + y + z = 1$ will be a plane in three dimensions that contains these three points, as shown in Figure 1.3. Hence when we consider a linear system in three unknowns, we are looking for a point in the three dimensional space that lies on all the planes described by the equations.

Activity 1.3. In this activity we examine the geometry of linear systems of three equations in three unknowns. Recall that each linear equation in three variables has a plane as its solution set. Use a piece of paper to represent each plane.

- (a) Is it possible for a general system of three linear equations in three unknowns to have no solutions? If so, geometrically describe this situation and then illustrate each with a specific example using a system of equations. If not, explain why not.
- (b) Is it possible for a general system of three linear equations in three unknowns to have exactly one solution? If so, geometrically describe this situation and then illustrate each with a specific example using a system of equations. If not, explain why not.
- (c) Is it possible for a general system of three linear equations in three unknowns to have infinitely many solutions? If so, geometrically describe this situation and then illustrate each with a specific example using a system of equations. If not, explain why not.



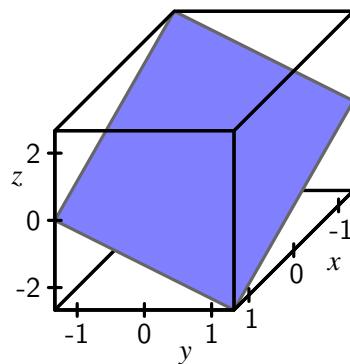


Figure 1.3: The plane $x + y + z = 1$.

Examples

What follows are worked examples that use the concepts from this section.

Example 1.5. Apply the allowable operations on equations to solve the system

$$\begin{aligned}x_1 + 2x_2 + x_3 - x_4 &= 4 \\-x_2 - x_3 + 3x_4 &= 6 \\x_1 + 2x_3 - x_4 &= 1 \\2x_1 - 3x_2 + x_3 + x_4 &= 2.\end{aligned}$$

Example Solution. We begin by eliminating the variable x_1 from all but the first equation. To do so, we replace the third equation with the third equation minus the first equation to obtain the equivalent system

$$\begin{aligned}x_1 + 2x_2 + x_3 - x_4 &= 4 \\-x_2 - x_3 + 3x_4 &= 6 \\-2x_2 + x_3 &= -3 \\2x_1 - 3x_2 + x_3 + x_4 &= 2.\end{aligned}$$

Then we replace the fourth equation with the fourth equation minus 2 times the first to obtain the equivalent system

$$\begin{aligned}x_1 + 2x_2 + x_3 - x_4 &= 4 \\-x_2 - x_3 + 3x_4 &= 6 \\-2x_2 + x_3 &= -3 \\-7x_2 - x_3 + 3x_4 &= -6.\end{aligned}$$

To continue the elimination process, we want to eliminate the x_2 variable from our latest third and fourth equations. To do so, we use the second equation so that we do not reinstate an x_1

variable in our new equations. We replace equation three with equation 3 minus 2 times equation 2 to produce the equivalent system

$$\begin{aligned}x_1 + 2x_2 + x_3 - x_4 &= 4 \\- x_2 - x_3 + 3x_4 &= 6 \\3x_3 - 6x_4 &= -15 \\- 7x_2 - x_3 + 3x_4 &= -6.\end{aligned}$$

Then we replace equation four with equation four minus 7 times equation 2, giving us the equivalent system

$$\begin{aligned}x_1 + 2x_2 + x_3 - x_4 &= 4 \\- x_2 - x_3 + 3x_4 &= 6 \\3x_3 - 6x_4 &= -15 \\6x_3 - 18x_4 &= -48.\end{aligned}$$

With one more step we can determine the value of x_4 . We use the last two equations to eliminate x_3 from the fourth equation by replacing equation four with equation four minus 2 times equation 3. This results in the equivalent system

$$\begin{aligned}x_1 + 2x_2 + x_3 - x_4 &= 4 \\- x_2 - x_3 + 3x_4 &= 6 \\3x_3 - 6x_4 &= -15 \\-6x_4 &= -18.\end{aligned}$$

The last equation tells us that $-6x_4 = -18$, or $x_4 = 3$. Substituting into the third equation shows that

$$\begin{aligned}3x_3 - 6(3) &= -15 \\3x_3 &= 3 \\x_3 &= 1.\end{aligned}$$

The second equation shows that

$$\begin{aligned}-x_2 - 1 + 3(3) &= 6 \\-x_2 &= -2 \\x_2 &= 2.\end{aligned}$$

Finally, the first equation tells us that

$$\begin{aligned}x_1 + 2(2) + 1 - 3 &= 4 \\x_1 &= 2.\end{aligned}$$

So the solution to our system is $x_1 = 2$, $x_2 = 2$, $x_3 = 1$, and $x_4 = 3$. It is worth substituting back into our original system to check to make sure that we have not made any arithmetic mistakes.



Example 1.6. A mining company has three mines. One day of operation at the mines produces the following output.

- Mine 1 produces 25 tons of copper, 600 kilograms of silver and 15 tons of manganese.
- Mine 2 produces 30 tons of copper, 500 kilograms of silver and 10 tons of manganese.
- Mine 3 produces 20 tons of copper, 550 kilograms of silver and 12 tons of manganese.

Suppose the company has orders for 550 tons of copper, 11350 kilograms of silver and 250 tons of manganese.

Write a system of equations to answer the question: how many days should the company operate each mine to exactly fill the orders? State clearly what the variables in your system represent. Then find the general solution of your system.

Example Solution. For our system, let x_1 be the number of days mine 1 operates, x_2 be the number of days mine 2 operates, and x_3 be the number of days mine 3 operates. Since mine 1 produces 25 tons of copper each day, in x_1 days mine 1 will produce $25x_1$ tons of copper. Mine 2 produces 30 tons of copper each day, so in x_2 days mine 2 will produce $30x_2$ tons of copper. Also, mine 3 produces 20 tons of copper each day, so in x_3 days mine 3 will produce $20x_3$ tons of copper. Since the company needs to supply a total of 550 tons of copper, we need to have $25x_1 + 30x_2 + 20x_3 = 550$. Similar analyses of silver and manganese give us the system

$$\begin{aligned} 25x_1 + 30x_2 + 20x_3 &= 550 \\ 600x_1 + 500x_2 + 550x_3 &= 11350 \\ 15x_1 + 10x_2 + 12x_3 &= 250. \end{aligned}$$

To solve the system, we eliminate the variable x_2 from the second and third equations by replacing equation two with equation two minus 24 times equation one and replacing equation three with equation three minus $\frac{3}{5}$ times equation one. This produces the equivalent system

$$\begin{aligned} 25x_1 + 30x_2 + 20x_3 &= 550 \\ -220x_2 + 70x_3 &= -1850 \\ -8x_2 &= -80 . \end{aligned}$$

We are fortunate now that we can determine the value of x_2 from the third equation, which tells us that $x_2 = 10$. Substituting into the second equation shows that

$$\begin{aligned} -220(10) + 70x_3 &= -1850 \\ 70x_3 &= 350 \\ x_3 &= 5. \end{aligned}$$

Substituting into the first equation allows us to determine the value for x_1 :

$$\begin{aligned} 25x_1 + 30(10) + 20(5) &= 550 \\ 25x_1 &= 150 \\ x_1 &= 6. \end{aligned}$$

So the company should run mine 1 for 6 days, mine 2 for 10 days, and mine 3 for 5 days to meet this demand.



Summary

In this section we introduced linear equations and systems of linear equations.

- Informally, a linear equation is an equation in which each term is either a constant or a constant times a variable. More formally, a linear equation in the variables x_1, x_2, \dots, x_n is an equation of the form

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n = b,$$

where n is a positive integer and a_1, a_2, \dots, a_n and b are constants.

- A system of linear equations is a collection of one or more linear equations in the same variables.
- Informally, a solution to a system of linear equations is a point that satisfies all of the equations in the system. More formally, a solution to a system of linear equation in n variables x_1, x_2, \dots, x_n is an ordered n -tuple (s_1, s_2, \dots, s_n) of numbers so that we obtain all true statements in the system when we replace x_1 with s_1 , x_2 with s_2, \dots , and x_n with s_n .
- Two linear systems are equivalent if they have the same solution set.
- The following operations on a system of equations do not change the solution set:
 - (1) Replace one equation by the sum of that equation and a scalar multiple of another equation.
 - (2) Interchange two equations.
 - (3) Replace an equation by a nonzero scalar multiple of itself.

Exercises

- (1) In the method of elimination there are three operations we can apply to solve a system of linear equations. For this exercise we focus on a system of equations in three unknowns x_1, x_2 , and x_3 , but the arguments generalize to a system with any number of variables. Consider the general system of three equations in three unknowns

$$\begin{aligned} 4x_1 - 4x_2 + 4x_3 &= 0 \\ 4x_1 + 2x_2 &= 8 \\ 2x_2 + 5x_3 &= 9. \end{aligned}$$

The goal of this exercise is to understand why the three operations on a system do not change the solutions to the system. Recall that a solution to a system with unknowns x_1, x_2 , and x_3 is a set of three numbers, one for x_1 , one for x_2 , and one for x_3 that satisfy all of the equations in the system.

- (a) Explain why, if we have a solution to this system, then that solution is also a solution to any constant k times the second equation.
- (b) Explain why, if we have a solution to this system, then that solution is also a solution to the sum of the first equation and k times the third equation for any constant k .



- (2) Alice stopped by a coffee shop two days in a row at a conference to buy drinks and pastries. On the first day, she bought a cup of coffee and two muffins for which she paid \$6.87. The next day she bought two cups of coffee and three muffins (for herself and a friend). Her bill was \$11.25. Use the method of elimination to determine the price of a cup of coffee, and the price of a muffin. Clearly explain your set-up for the problem (Assume you are explaining your solution to someone who has not solved the problem herself/himself).
- (3) Alice stopped by a coffee shop three days in a row at a conference to buy drinks and pastries. On the first day, she bought a cup of coffee, a muffin and a scone for which she paid \$6.15. The next day she bought two cups of coffee, three muffins and a scone (for herself and friends). Her bill was \$12.20. The last day she bought a cup of coffee, two muffins and two scones, and paid \$10.35. Determine the price of a cup of coffee, the price of a muffin and the price of a scone. Clearly explain your set-up for the problem (Assume you are explaining your solution to someone who has not solved the problem herself/himself).
- (4) (a) Find an example of a system of two linear equations in variables x, y for each of the following three cases:
- where the equations correspond to two non-parallel lines,
 - two parallel distinct lines,
 - two identical lines (represented with different equations).
- (b) Describe how the relationship between the coefficients of the variables of the two equations in parts (ii) and (iii) are different than the relationship between those coefficients in part (i) (Note: Please make sure your system examples are different than the examples in the activities, and that they are your own examples.)
- (5) In a grid of wires in thermal equilibrium, the temperature at interior nodes is the average of the temperatures at adjacent nodes. Consider the grid as shown in Figure 1.4, with x_1, x_2 , and x_3 the temperatures (in degrees Centigrade) at the indicated interior nodes, and fixed temperatures at the other nodes as shown. For example, the nodes adjacent to the node with temperature x_1 have temperatures of $x_2, 200, 0$, and 0 , so when the grid is in thermal equilibrium x_1 is the average of these temperatures:

$$x_1 = \frac{x_2 + 200 + 0 + 0}{4}.$$

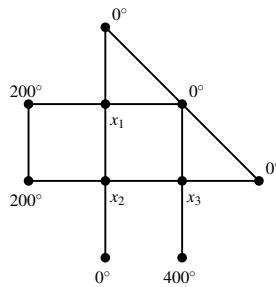


Figure 1.4: A grid of wires.

- (a) Determine equations for the temperatures x_2 and x_3 if the grid is in thermal equilibrium to construct a system of three linear equations in x_1 , x_2 , and x_3 that models node temperatures in the grid in thermal equilibrium.
- (b) Use the method of elimination to find a specific solution to the system that makes sense in context.
- (6) We have seen that a linear system of two equations in two unknowns can have no solutions, one solution, or infinitely many solutions. Find, if possible, a specific example of each of the following. If not possible, explain why.
- A linear system of three equations in two unknowns with no solutions.
 - A linear system of three equations in two unknowns with exactly one solution.
 - A linear system of three equations in two unknowns with exactly two solutions.
 - A linear system of three equations in two unknowns with infinitely many solutions.
- (7) We have seen that a linear system of three equations in three unknowns can have no solutions, one solution, or infinitely many solutions. Find, if possible, a specific example of each of the following. If not possible, explain why.
- A linear system of two equations in three unknowns with no solutions.
 - A linear system of two equations in three unknowns with exactly one solution.
 - A linear system of two equations in three unknowns with exactly two solutions.
 - A linear system of two equations in three unknowns with infinitely many solutions.
- (8) Find a system of three linear equations in two variables u , v whose solution is $u = 2$, $v = 1$.
- (9) Consider the system of linear equations

$$\begin{aligned}x_1 + hx_2 &= 2 \\3x_1 + 5x_2 &= 1\end{aligned}$$

where h is an unknown constant.

- Determine the solution(s) of this system for all possible h values, if a solution exists.
(Note: Your answers for the variables will depend on the h .)
- How do your answers change if the second equation in the system above is changed to $3x_1 + 5x_2 = 6$?

- (10) Suppose we are given a system of two linear equations

$$x_1 + 2x_2 - x_3 = 1 \quad (1.7)$$

$$3x_1 + x_2 + 2x_3 = -1. \quad (1.8)$$

Find another system of two linear equations E_1 and E_2 in the variables x_1 , x_2 , and x_3 that are not multiples of each other or of equations (1.7) or (1.8) so that any solution (x_1, x_2, x_3) to the system (1.7) and (1.8) is a solution to the system E_1 and E_2 .



True/False Questions

In many sections you will be given True/False questions. In each of the True/False questions, you will be given a statement, such as “If we add corresponding sides of two linear equations, then the resulting equation is a linear equation.” and “One can find a system of two equations in two unknowns that has infinitely many solutions.”. Your task will be to determine the truth value of the statement and to give a brief justification for your choice.

Note that a *general* statement is considered *true* only when it is always true. For example, the first of the above statements, “If we add corresponding sides of two linear equations, then the resulting equation is a linear equation.”, is a general statement. For this statement to be true, the equation we obtain by adding corresponding sides of any two linear equations has to be linear. If we can find two equations that do not give a linear equation when combined in this way, then this statement is false.

Note that an *existential* statement is considered *true* if there is at least one example which makes it true. For example, the latter of the above statements, “One can find a system of two equations in two unknowns that has infinitely many solutions.”, is an existential statement. For this statement to be true, existence of a system of two equations in two unknowns with infinitely many solutions should suffice. If it is impossible to find two such equations, then this statement is false.

To justify that something always happens or never happens, one would need to refer to other statements whose truth is known, such as theorems, definitions. In particular, giving an **example** of two linear equations that produce a linear equation when we add corresponding sides *does not justify* why the sum of **any** two linear equations is also linear. Using the definition of linear equations, however, we can justify why this new equation will always be linear: each side of a linear equation is linear, and adding linear expressions always produces a linear sum.

To justify that there are examples of something happening or not happening, one would need to give a specific example. For example, in justifying the claim that there is a system of two equations in two unknowns with infinitely many solutions, it is not enough to say “An equation in two unknowns is a line in the xy -plane, so there can be two equations with the same line as their solution.”. In general, you should avoid the words “can”, “possibly”, “maybe”, etc., in your justifications. Instead, giving an example such as “The linear system $x + y = 1$ and $2x + 2y = 2$ of two equations in two unknowns has infinitely many solutions since the second equation gives the same line as the first in the xy -plane.” provides complete justification beyond a reasonable doubt.

Each response to a True/False statement should be more than just True or False. It is important that you provide **justification** for your responses.

- (1) (a) **True/False** The set of all solutions of a linear equation can be represented graphically as a line.
- (b) **True/False** The set of all solutions of a linear equation in two variables can be represented graphically as a line.
- (c) **True/False** The set of all solutions of an equation in two variables can be represented graphically as a line.
- (d) **True/False** A system of three linear equations in two unknowns cannot have a unique solution.



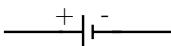
- (e) **True/False** A system of three linear equations in three unknowns has a unique solution.

Project: Modeling an Electrical Circuit and the Wheatstone Bridge Circuit

Mathematical modeling, or the act of creating equations to model given information, is an important part of problem solving. In this section we will see how we derived the system of equations

$$\begin{aligned} I_1 - I_2 + I_3 &= 0 \\ 5I_1 + 2I_2 &= 8 \\ 2I_2 + 4I_3 &= 5, \end{aligned}$$

to represent the electrical current in the circuit shown in Figure 1.2. Recall that a circuit consists of

- one or more electrical sources (like a battery), denoted by 
- one or more resistors (like any appliance that you plug into a wall outlet), denoted by .

The source creates a charge that produces potential energy E measured in volts (V). No substance conducts electricity perfectly, there is always some price to pay (energy loss) to moving electricity. Electrical current I in amperes (A) is the flow of the electric charge in the circuit. (A current of 1 ampere means that 6.2×10^{18} electrons pass through the circuit per second.) Current flows out of the positive terminal of a source and runs through each branch of the circuit. Let I_1 be the current flowing through the upper branch, I_2 the current through middle branch, and I_3 the current through the lower branch as illustrated in Figure 1.2. The goal is to find the current flowing in each branch of the circuit.

Linear algebra comes into play when analyzing a circuit based on the relationship between current, resistance, and potential. Three basic principles govern current low in a circuit.

- (1) Resistance R in ohms (Ω) can be thought of as a measure of how difficult it is to move a charge along a circuit. When a current flows through a resistor, it must expend some energy, called a *voltage drop*. Ohm's Law states that the voltage drop E across a resistor is the product of the current I passing through the resistor and the resistance R . That is,

$$E = IR.$$

- (2) Kirchoff's Current Law states that at any point in an electrical circuit, the sum of currents flowing into that point is equal to the sum of currents flowing out of that point.
- (3) Kirchoff's Voltage Law says that around any closed loop the sum of the voltage drops is equal to the sum of the voltage rises.

To see how these laws allow us to model the circuit in Figure 1.2, we will need three equations in I_1 , I_2 , and I_3 to determine the values of these currents. Let us first apply Kirchoff's Current Law to the point P. The currents flowing into point P are I_1 and I_3 , and the current flowing out is I_2 . This produces the equation $I_1 + I_3 = I_2$, or

$$I_1 - I_2 + I_3 = 0.$$

Project Activity 1.1. Apply Kirchoff's Current Law to the point Q to obtain an equation in I_1 , I_2 , and I_3 . What do you notice?

We have three variables to determine, so we still need two more equations in I_1 , I_2 , and I_3 . Next we apply Kirchoff's Voltage Law to the top loop in the circuit in Figure 1.2. We will assume the following sign conventions:

- A current passing through a resistor produces a voltage drop if it flows in the direction of loop (and a voltage rise if the current passes in the opposite direction of the loop).
- A current passing through a source in the direction of the loop produces a voltage drop if it flows from + to – and a voltage rise if it flows from – to +, while a current passing through a source in the opposite direction of the loop produces a voltage rise if it flows from + to – and a voltage drop if it flows from – to +.

(The directions chosen in Figure 1.2 for the voltage flow are arbitrary – if we reverse the flow then we just replace voltage drops with voltage rises and obtain the same equations. If a solution shows that a current is negative, then that current flows in the direction opposite of what is shown.)

If we move in the counterclockwise direction around the top loop in the circuit in Figure 1.2, there is a voltage rise through the source of 8 volts. This must equal the voltage drop in this loop. The current I_1 passing through the resistor of resistance 2Ω produces a voltage drop of $2I_1$ volts. Similarly, the current I_1 passing through the resistor of resistance 3Ω produces a voltage drop of $3I_1$ volts. Finally, the current I_2 passing through the resistor of resistance 2Ω produces a voltage drop of $2I_2$ volts. So Kirchoff's Voltage Law applied to the top loop in the circuit in Figure 1.2 gives us the equation $2I_1 + 3I_1 + 2I_2 = 8$ or

$$5I_1 + 2I_2 = 8.$$

Project Activity 1.2. Apply Kirchoff's Voltage Law to the bottom loop in the circuit in Figure 1.2 to obtain an equation in I_1 , I_2 , and I_3 . Compare the three equations we have found to those in the introduction.

Project Activity 1.3. Consider the circuit as shown in Figure 1.5, with a single source and five resistors with resistances R_1 , R_2 , R_3 , R_4 , and R_5 as labeled.

- (a) Assume the following information. The voltage E is 13 volts, $R_1 = R_2 = R_3 = R_5 = 1\Omega$, and $R_4 = 2\Omega$. Follow the directions given to find the currents I_0 , I_1 , I_2 , I_3 , I_4 , and I_5 .
- Use Kirchoff's Current Law to show that $I_0 = I_1 + I_2$, $I_3 = I_1 - I_5$, and $I_4 = I_2 + I_5$. Thus, we reduce the problem to three variables.



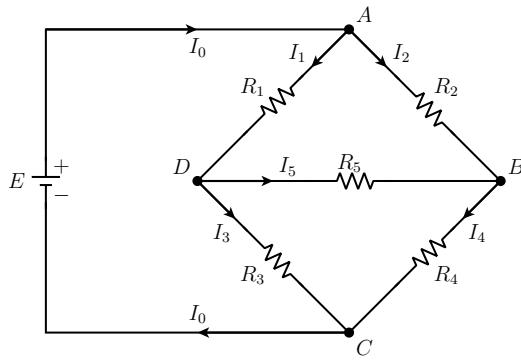


Figure 1.5: A Wheatstone bridge circuit.

- ii. Apply Kirchoff's Voltage Law to three loops to show that the currents must satisfy the linear system

$$2I_1 - I_5 = 13 \quad (1.9)$$

$$3I_2 + 2I_5 = 13 \quad (1.10)$$

$$I_1 - I_2 + I_5 = 0. \quad (1.11)$$

- iii. Solve the system to find the unknown currents.

- (b) The circuit pictured in Figure 1.5 is called a *Wheatstone bridge* (invented by Samuel Hunter Christie in 1833 and popularized by Sir Charles Wheatstone in 1843). The Wheatstone bridge is a circuit designed to determine an unknown resistance by balancing two paths in a circuit. It is set up so that the resistances of resistors R_1 and R_2 are known, R_3 is a variable resistor and we want to find the resistance of R_4 . The resistor R_5 is replaced with a voltmeter, and the resistance of R_3 is varied until the voltmeter reads 0. This balances the circuit and tells the resistance of resistor R_4 . Show that if the current I_5 in Figure 1.5 is 0 (so the circuit is balanced), then $R_4 = \frac{R_2 R_3}{R_1}$ (which is how we calculate the unknown resistance R_4). Do this in general and do not use any specific values for the resistances or the voltage.

Section 2

The Matrix Representation of a Linear System

Focus Questions

By the end of this section, you should be able to give precise and thorough answers to the questions listed below. You may want to keep these questions in mind to focus your thoughts as you complete the section.

- What is a matrix?
- How do we associate a matrix to a system of linear equations?
- What row operations can we perform on an augmented matrix of a linear system to solve the system of linear equations?
- What are pivots, basic variables, and free variables?
- How many solutions can a system of linear equations have?
- When is a linear system consistent?
- When does a linear system have infinitely many solutions? A unique solution?
- How can we represent the set of solutions to a consistent system if the system has infinitely many solutions?

Application: Simpson's Rule

You may recall that Simpson's Rule from calculus ($\frac{2}{3}$ of the midpoint approximation plus $\frac{1}{3}$ of the trapezoid approximation) is a formula that can be used to approximate definite integrals. On the one hand, Simpson's Rule is a weighted average of the midpoint and trapezoid sum, but that does not completely explain why Simpson's Rule is so much better than either the midpoint or trapezoid

sum. While the midpoint and trapezoid sums use line segments to approximate a function on an interval, Simpson's Rule uses parabolas. In order to use Simpson's Rule, we need to know how to exactly fit a quadratic function to three points. More details about this process can be found at the end of this section. This idea of fitting a polynomial to a set of data points has uses in other areas as well. For example, two common applications of Bézier curves are font design and drawing tools. When fitting a polynomial to a large set of data points, our systems of equations can become quite large, and can be difficult to solve by hand. In this section we will see how to use matrices to more conveniently represent systems of equations of any size. We also consider how the elimination process works on the matrix representation of a linear system and how we can determine the existence of solutions and the form of solutions of a linear system.

Introduction

When working with a linear system, the labels for the variables are irrelevant to the solution – the only thing that matters is the coefficients of the variables in the equations and the constants on the other side of the equations. For example, given a linear system of the form

$$\begin{aligned} a_2 - a_1 + a_0 &= 2 \\ a_2 + a_1 + a_0 &= 6 \\ 4a_2 + 2a_1 + a_0 &= 5, \end{aligned} \tag{2.1}$$

the important information in the system can be represented as

$$\begin{array}{cccc} 1 & -1 & 1 & 2 \\ 1 & 1 & 1 & 6 \\ 4 & 2 & 1 & 5 \end{array}$$

where we interpret the first three numbers in each horizontal row to represent the coefficients of the variables a , b and c , respectively, and the last number to be the constant on the right hand side of the equation. This tells us that we can record all the necessary information about our system in a rectangular array of numbers. Such an array is called a *matrix*.

Definition 2.1. A **matrix** is a rectangular array of quantities or expressions.

We usually delineate a matrix by enclosing its entries in square brackets $[*]$. For the system in (2.1), there are two corresponding matrices:

$$\left[\begin{array}{ccc|c} 1 & -1 & 1 & 2 \\ 1 & 1 & 1 & 6 \\ 4 & 2 & 1 & 5 \end{array} \right]$$

The matrix on the left is the matrix of the coefficients of the system, and is called the *coefficient matrix* of the system. The matrix on the right is the matrix of coefficients and the constants, and is called the *augmented matrix* of the system (where we say we augment the coefficient matrix with the additional column of constants). We will separate the augmented column from the coefficient



matrix with a vertical line to keep it clear that the last column is an augmented column of constants and not a column of coefficients.¹

Terminology. There is some important terminology related to matrices.

- Any number in a matrix is called an *entry* of the matrix.
- The collection of entries in an augmented matrix that corresponds to a given equation (that is reading the entries from left to right, or a horizontal set of entries) is called a *row* of the matrix. We number the rows from top to bottom in a matrix. For example, $\begin{bmatrix} 1 & -1 & 1 \end{bmatrix}$ is the first row and $\begin{bmatrix} 1 & 1 & 1 \end{bmatrix}$ is the second row of the coefficient matrix of the system (2.1).
- The set of entries as we read from top to bottom (or a vertical set of entries that correspond to one fixed variable or the constants on the right hand sides of the equations) is called a *column* of the matrix. We number the columns from left to right in a matrix. For example, $\begin{bmatrix} 1 \\ 1 \\ 4 \end{bmatrix}$ is the first column and $\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ is the third column of the coefficient matrix of the system (2.1).
- The *size* of a matrix is given as $m \times n$ where m is the number of rows and n is the number of columns. The coefficient matrix above is a 3×3 matrix since it has 3 rows and 3 columns, while the augmented matrix is a 3×4 matrix as it has 4 columns.

Preview Activity 2.1.

- (1) Write the augmented matrix for the following linear system. If needed, rearrange an equation to ensure that the variables appear in the same order on the left side in each equation with the constants being on the right hand side of each equation.

$$\begin{aligned} -x_3 + 3 + 2x_2 &= -x_1 \\ -3 + 2x_3 &= -x_2 \\ -2x_2 + x_1 &= 3x_3 - 7 \end{aligned} \tag{2.2}$$

- (2) Write the linear system in variables x_1, x_2 and x_3 , appearing in the natural order that corresponds to the following augmented matrix. Then solve the linear system using the elimination method.

$$\left[\begin{array}{ccc|c} 1 & 1 & -1 & 4 \\ 1 & 2 & 2 & 3 \\ 2 & 3 & -3 & 11 \end{array} \right]$$

- (3) Consider the three types of elementary operations on systems of equations introduced in Section 1. Each row of an augmented matrix of a system corresponds to an equation, so each elementary operation on equations corresponds to an operation on rows (called row operations).

¹You should note that not every author uses this convention – when they do not, it is important that you be careful to understand if the matrix has an augmented column or not.

- (a) Describe the row operation that corresponds to interchanging two equations.
- (b) Describe the row operation that corresponds to multiplying an equation by a nonzero scalar.
- (c) Describe the row operation that corresponds to replacing one equation by the sum of that equation and a scalar multiple of another equation.

Simplifying Linear Systems Represented in Matrix Form

Once we have stored the information about a linear system in an augmented matrix, we can perform the elementary operations directly on the augmented matrix.

Recall that the allowable operations on a system of equations are the following:

- (1) Replacing one equation by the sum of that equation and a scalar multiple of another equation.
- (2) Interchanging the positions of two equations.
- (3) Replacing an equation by a nonzero scalar multiple of itself.

Recall that we use these elementary operations to transform a system, with the ultimate goal of finding a simpler, equivalent system that we can solve. Since each row of an augmented matrix corresponds to an equation, we can translate these operations on equations to corresponding operations on rows (called *row operations* or *elementary row operations*):

- (1) Replacing one row by the sum of that row and a scalar multiple of another row.
- (2) Interchanging two rows.
- (3) Replacing a row by a nonzero scalar multiple of itself.

Activity 2.1. Consider the system

$$\begin{aligned} a_2 - a_1 + a_0 &= 2 \\ a_2 + a_1 + a_0 &= 6 \\ 4a_2 + 2a_1 + a_0 &= 5 \end{aligned}$$

with corresponding augmented matrix

$$\left[\begin{array}{ccc|c} 1 & -1 & 1 & 2 \\ 1 & 1 & 1 & 6 \\ 4 & 2 & 1 & 5 \end{array} \right]$$

- (a) As a first step in solving our system, we might eliminate a_2 from the second equation. This means that the corresponding entry in the second row and first column of the augmented matrix will become 0. Find a row operation that adds a multiple of the first row to the second row to achieve this goal. Then write the system of equations that corresponds to this new augmented matrix.



- (b) Now that we have eliminated the a_2 terms from the second equation, we eliminate the a_2 term from the third equation. Find an appropriate row operation that does that, and write the corresponding system of linear equations that corresponds to the new augmented matrix.
- (c) Now you should have a system in which the last two rows correspond to a system of 2 linear equations in two unknowns. Use a row operation that adds a multiple of the second row to the third row to turn the coefficient of a_1 in the third row to 0. Then write the corresponding system of linear equations.
- (d) Your simplified system and its augment matrix are in *row echelon form* and this system is solvable using *back-substitution* (substituting the known variable values into the previous equation to find the value of another variable). Solve the system.

Reflection 1. Do you see how this standard elimination process can be generalized to any linear system with any number of variables to produce a simplified system? Do you see why the process does not change the solutions of the system? If needed, can you modify the standard elimination process to obtain a simplified system in which the last equation contains only the variable a_2 , the next to last equation contains only the variables a_1, a_2 , etc.? Understanding the standard process will enable you to be able to modify it, if needed, in a problem.

Activity 2.1 illustrates how we can perform all of the operations on equations with operations on the rows of augmented matrices to reduce a system to a solvable form. Each time we perform an operation on the system of equations (or on the rows of an augmented matrix) we obtain an equivalent system (or an augmented matrix corresponding to an equivalent system). For completeness, we list the operations on equations and the corresponding row operations below that can be used to solve our polynomial fitting system. Throughout the process we will let E_1 , E_2 , and E_3 be the first, second, and third equations in the system and R_1 , R_2 , and R_3 the first, second, and third rows of the augmented matrices. The notation $E_1 + E_2$ placed next to equation E_2 means that we replace the second equation in the system with the sum of the first two equations. We start with the system

$$\begin{aligned} a_2 - a_1 + a_0 &= 2 \\ a_2 + a_1 + a_0 &= 6 \\ 4a_2 + 2a_1 + a_0 &= 5 \end{aligned}$$

On the left we demonstrate the operations on equations and on the right the corresponding operations on rows of the augmented matrix.

$$\begin{array}{ll} a_2 - a_1 + a_0 = 2 & \\ 2a_1 = 4 & \\ 4a_2 + 2a_1 + a_0 = 5 & \end{array} \quad \begin{array}{l} R_2 - R_1 \rightarrow R_2 \\ \hline \end{array} \quad \left[\begin{array}{ccc|c} 1 & -1 & 1 & 2 \\ 0 & 2 & 0 & 4 \\ 4 & 2 & 1 & 5 \end{array} \right]$$

$$\begin{array}{ll} a_2 - a_1 + a_0 = 2 & \\ 2a_1 = 4 & \\ 6a_1 - 3a_0 = -3 & \end{array} \quad \begin{array}{l} R_3 - 4R_1 \rightarrow R_3 \\ R_3 - 4R_1 \rightarrow R_3 \\ \hline \end{array} \quad \left[\begin{array}{ccc|c} 1 & -1 & 1 & 2 \\ 0 & 2 & 0 & 4 \\ 0 & 6 & -3 & -3 \end{array} \right]$$



$$\begin{array}{rcl}
 a_2 - a_1 + a_0 & = & 2 \\
 2a_1 & = & 4 \\
 -3a_0 & = & -15
 \end{array}
 \quad
 \left[\begin{array}{ccc|c}
 1 & -1 & 1 & 2 \\
 0 & 2 & 0 & 4 \\
 0 & 0 & -3 & -15
 \end{array} \right]$$

$E_3 - 3E_2 \rightarrow E_3$ $R_3 - 3R_2 \rightarrow R_3$

Now we can solve the last equation for a_0 to find that $a_0 = 5$. The second equation gives us $a_1 = 2$.² Finally, using the first equation with the already determined values of a_0 and a_1 gives us $a_2 = -1$. Thus we have found the solution to the polynomial fitting system to be $a_2 = -1$, $a_1 = 2$, and $a_0 = 5$.

We summarize the steps of the (partial) elimination on matrices we used above to solve a general linear system in the variables x_1, x_2, \dots, x_n .

- (1) Interchange equations if needed to ensure that the coefficient of x_1 (or, more generally, the first non-zero variable) in the first equation is non-zero.
- (2) Use the first equation to eliminate x_1 (or, the first non-zero variable) from other equations by adding a multiple of the first equation to the others.
- (3) After x_1 is eliminated from all equations but the first equation, focus on the rest of the equations. Repeat the process of elimination on these equations to eliminate x_2 (or, the next non-zero variable) all but the second equation.
- (4) Once the process of eliminating variables recursively is finished, solve for the variables in a backwards fashion starting with the last equation and substituting known values in the equations above as they become known.

This elimination method where the variables are eliminated from lower equations is called the *forward elimination phase* as it eliminates variables in the forward direction. Solving for variables using substitution into upper equations is called *back substitution*. The matrix representation of a linear system after the forward elimination process is said to be in *row echelon form*. We will define this form and the elimination process on the matrices more precisely in the next section.

Linear Systems with Infinitely Many Solutions

Each of the systems that we solved so far have had a unique (exactly one) solution. The geometric representation of linear systems with two equations in two variables shows that this does not always have to be the case. We also have linear systems with no solution and systems with infinitely many solutions. We now consider the problem of how to represent the set of solutions of a linear system that has infinitely many solutions. (Systems with infinitely many solutions will also be of special interest to us a bit later when we study eigenspaces of a matrix.)

Activity 2.2. Consider the system

$$\begin{aligned}
 x_1 + 2x_2 - x_3 &= 1 \\
 x_1 + x_2 - 3x_3 &= 0 \\
 2x_1 + 3x_2 - 4x_3 &= 1.
 \end{aligned}$$

²If there had been an a_0 term in the second equation, we could have substituted $a_0 = 5$ and solved for a_1



- (a) Without explicitly solving the system, check that $(-1, 1, 0)$ and $(4, -1, 1)$ are solutions to this system.
- (b) Without explicitly solving the system, show that $x_1 = -1 + 5t$, $x_2 = 1 - 2t$, and $x_3 = t$ is a solution to this system for any value of t . What values of t yield the solutions $(-1, 1, 0)$ and $(4, -1, 1)$ from part (a)? The equations $x_1 = -1 + 5t$, $x_2 = 1 - 2t$, and $x_3 = t$ form what is called a *parametric solution* to the system with *parameter* t .
- (c) Part (b) shows that our system has infinitely many solutions. We were given solutions in part (b) – but how do we find these solutions and how do we know that these are all of the solutions? We address those questions now.

If we apply row operations to the augmented matrix

$$\left[\begin{array}{ccc|c} 1 & 2 & -1 & 1 \\ 1 & 1 & -3 & 0 \\ 2 & 3 & -4 & 1 \end{array} \right]$$

of this system, we can reduce this system to one with augmented matrix

$$\left[\begin{array}{ccc|c} 1 & 2 & -1 & 1 \\ 0 & 1 & 2 & 1 \\ 0 & 0 & 0 & 0 \end{array} \right].$$

- What is it about this reduced form of the augmented matrix that indicates that the system has infinitely many solutions?
- Since the system has infinitely many solutions, we will not be able to explicitly determine values for each of the variables. Instead, at least one of the variables can be chosen arbitrarily. What is it about the reduced form of the augmented matrix that indicates that x_3 is convenient to choose as the arbitrary variable?
- Letting x_3 be arbitrary (we call x_3 a *free variable*), use the second row to show that $x_2 = 1 - 2x_3$ (so that we can write x_2 in terms of the arbitrary variable x_3).
- Use the first row to show that $x_1 = 5x_3 - 1$ (and we can write x_1 in terms of the arbitrary variable x_3). Compare this to the solutions from part (b).

After using the elimination method, the first non-zero coefficient (from the left) of each equation in the linear system is in a different position. We call each such coefficient a *pivot* and a variable corresponding to a pivot a *basic variable*. In the system

$$\begin{aligned} a_2 - a_1 + a_0 &= 2 \\ 2a_1 &= 4 \\ -3a_0 &= -15 \end{aligned}$$

the basic variables are a_2, a_1, a_0 for the first, second, and third equations, respectively. In the system,



$$\begin{aligned}x_1 + 2x_2 - x_3 &= 1 \\x_2 + 2x_3 &= 1 \\0 &= 0\end{aligned}$$

the basic variables are x_1 and x_2 for the first and second equations, respectively, while the third equation does not have a basic variable. Through back-substitution, we can solve for each variable in a unique way if each appears as the basic variable in an equation. If, however, a variable is *free*, meaning that it is not the basic variable of an equation, we cannot solve for that variable explicitly. We instead assign a distinct parameter to each such free variable and solve for the basic variables in terms of these parameters.

Definition 2.2. The first non-zero coefficient (from the left) in an equation in a linear system after elimination is called a **pivot**. A variable corresponding to a pivot is a **basic variable** and while a variable not corresponding to a pivot is a **free variable**.

Activity 2.3. Each matrix is an augmented matrix for a linear system after elimination. Identify the basic variables (if any) and free variables (if any). Then write the general solution (if there is a solution) expressing all variables in terms of the free variables. Use any symbols you like for the variables.

$$(a) \left[\begin{array}{ccc|c} 1 & 0 & 2 & 1 \\ 0 & 3 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{array} \right]$$

$$(b) \left[\begin{array}{ccc|c} 1 & 0 & -1 & 1 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 \end{array} \right]$$

$$(c) \left[\begin{array}{cccc|c} 1 & 2 & -1 & 1 & 1 \\ 0 & 1 & 0 & 2 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array} \right]$$

Reflection 2. Does the existence of a row of 0's always mean a free variable? Can you think of an example where there is a row of 0's but none of the variables is free? How do the numbers of equations and the variables compare in that case?

Linear Systems with No Solutions

We saw in the previous section that geometrically two parallel and distinct lines represent a linear system with two equations in two unknowns which has no solution. Similarly, two parallel and distinct planes in three dimensions represent a linear system with two equations in three unknowns which has no solution. We can have at least four different geometric configurations of three planes in three dimensions representing a system with no solution. But how do these geometrical configurations manifest themselves algebraically?



Activity 2.4. Consider the linear system

$$\begin{aligned}x_1 - x_2 + x_3 &= 2 \\x_1 + x_2 - 3x_3 &= 1 \\3x_1 - x_2 - x_3 &= 6.\end{aligned}$$

- (a) Apply the elimination process to the augmented matrix of this system. Write the system of equations that corresponds to the final reduced matrix.
- (b) Discuss which feature in the final simplified system makes it easy to determine that the system has no solution. Similarly, what features in the matrix representation makes it easy to see the system has no solution?

We summarize our observations about when a system has a solution, and which of those cases has a unique solution.

Theorem 2.3. A linear system is consistent if after the elimination process there is no equation of the form $0 = b$ where b is a non-zero number. If a linear system is consistent and has a free variable, then it has infinitely many solutions. If it is consistent and has no free variables, then there is a unique solution.

Examples

What follows are worked examples that use the concepts from this section.

Example 2.4. Consider the linear system

$$\begin{aligned}x_1 - x_2 &\quad + 2x_4 = 1 \\2x_1 + 3x_2 - 2x_3 + 5x_4 &= 4 \\x_1 - x_2 + x_3 - x_4 &= 0 \\4x_1 + x_2 - x_3 + 6x_4 &= 5.\end{aligned}$$

- (a) Set up the augmented matrix for this linear system.
- (b) Find all solutions to the system using forward elimination.
- (c) Suppose, after forward elimination, the augmented matrix of the system

$$\begin{aligned}x_1 - x_2 &\quad + 2x_4 = 1 \\2x_1 + 3x_2 - 2x_3 + 5x_4 &= 4 \\x_1 - x_2 + x_3 - x_4 &= 0 \\4x_1 + x_2 - x_3 + 6x_4 &= h.\end{aligned}$$



has the form

$$\left[\begin{array}{cccc|c} 1 & -1 & 0 & 2 & 1 \\ 0 & 5 & -2 & 1 & 2 \\ 0 & 0 & 1 & -3 & -1 \\ 0 & 0 & 0 & 0 & h-5 \end{array} \right].$$

For which values of h does this system have:

- i. No solutions?
- ii. A unique solution? Find the solution.
- iii. Infinitely many solution? Determine all solutions?

Example Solution.

(a) The augmented matrix for this system is

$$\left[\begin{array}{cccc|c} 1 & -1 & 0 & 2 & 1 \\ 2 & 3 & -2 & 5 & 4 \\ 1 & -1 & 1 & -1 & 0 \\ 4 & 1 & -1 & 6 & 5 \end{array} \right].$$

(b) We apply forward elimination, first making the entries below the 1 in the upper left all 0. We do this by replacing row two with row two minus 2 times row 1, row three with row three minus row 1, and row four with row four minus 4 row one. This produces the augmented matrix

$$\left[\begin{array}{cccc|c} 1 & -1 & 0 & 2 & 1 \\ 0 & 5 & -2 & 1 & 2 \\ 0 & 0 & 1 & -3 & -1 \\ 0 & 5 & -1 & -2 & 1 \end{array} \right].$$

Now we eliminate the leading 5 in the fourth row by replacing row four with row four minus row two to obtain the augmented matrix

$$\left[\begin{array}{cccc|c} 1 & -1 & 0 & 2 & 1 \\ 0 & 5 & -2 & 1 & 2 \\ 0 & 0 & 1 & -3 & -1 \\ 0 & 0 & 1 & -3 & -1 \end{array} \right].$$

When we replace row four with row four minus row three, we wind up with a row of zeros:

$$\left[\begin{array}{cccc|c} 1 & -1 & 0 & 2 & 1 \\ 0 & 5 & -2 & 1 & 2 \\ 0 & 0 & 1 & -3 & -1 \\ 0 & 0 & 0 & 0 & 0 \end{array} \right].$$

We see that there is no pivot in column four, so x_4 is a free variable. We can solve for the other variables in terms of x_4 . The third row shows us that

$$\begin{aligned} x_3 - 3x_4 &= -1 \\ x_3 &= 3x_4 - 1. \end{aligned}$$



The second row tells us that

$$\begin{aligned} 5x_2 - 2x_3 + x_4 &= 2 \\ 5x_2 &= 2x_3 - x_4 + 2 \\ 5x_2 &= 2(3x_4 - 1) - x_4 + 2 \\ 5x_2 &= 5x_4 \\ x_2 &= x_4. \end{aligned}$$

Finally, the first row gives us

$$\begin{aligned} x_1 - x_2 + 2x_4 &= 1 \\ x_1 &= x_2 - 2x_4 + 1 \\ x_1 &= x_4 - 2x_4 + 1 \\ x_1 &= -x_4 + 1. \end{aligned}$$

So this system has infinitely many solutions, with $x_1 = -x_4 + 1$, $x_2 = x_4$, $x_3 = 3x_4 - 1$, and x_4 is arbitrary. As a check, notice that

$$(-x_4 + 1) - x_4 + 2x_4 = 1$$

and so this solution satisfies the first equation in our system. You should check to verify that it also satisfies the other three equations.

- (c)
 - i. The system has no solutions when there is an equation of the form $0 = b$ for some nonzero number b . The last row will correspond to an equation of the form $0 = h - 5$. So our system will have no solutions when $h \neq 5$.
 - ii. When $h \neq 5$, the system has no solutions. When $h = 5$, the variable x_4 is a free variable and the system has infinitely many solutions. So there are no values of h for which the system has exactly one solution.
 - iii. When $h = 5$, the variable x_4 is a free variable and the system has infinitely many solutions. The solutions were already found in part (a).

Example 2.5. After applying row operations to the augmented matrix of a system of linear equations, each of which describes a plane in 3-space, the following augmented matrix was obtained:

$$\left[\begin{array}{ccc|c} 1 & a & 0 & 2 \\ 0 & 2 - 2a & b & -4 \\ 0 & 0 & 3 - \frac{1}{2}b & 1 \end{array} \right].$$

- (a) Describe, algebraically and geometrically, all solutions (if any), to this system when $a = 0$ and $b = 2$.
- (b) Describe, algebraically and geometrically, all solutions (if any), to this system when $a = 0$ and $b = 6$.

- (c) Describe, algebraically and geometrically, all solutions (if any), to this system when $a = 1$ and $b = 12$.

Example Solution. Throughout, we will let the variables x , y , and z correspond to the first, second, and third columns, respectively, of our augmented matrix.

- (a) When $a = 0$ and $b = 2$ our augmented matrix has the form

$$\left[\begin{array}{ccc|c} 1 & 0 & 0 & 2 \\ 0 & 2 & 2 & -4 \\ 0 & 0 & 2 & 1 \end{array} \right].$$

This matrix corresponds to the system

$$\begin{aligned} x &= 2 \\ 2y + 2z &= -4 \\ 2z &= 1. \end{aligned}$$

There are no equations of the form $0 = b$ for a nonzero constant b , so the system is consistent. There are no free variables, so the system has a unique solution. Algebraically, the solution is $x = 2$, $z = \frac{1}{2}$, and $y = -\frac{5}{2}$. Geometrically, this tells us that the three planes given by the original system intersect in a single point.

- (b) When $a = 0$ and $b = 6$ our augmented matrix has the form

$$\left[\begin{array}{ccc|c} 1 & 0 & 0 & 2 \\ 0 & 2 & 6 & -4 \\ 0 & 0 & 0 & 1 \end{array} \right].$$

The last row corresponds to the equation $0 = 1$, so our system is inconsistent and has no solution. Geometrically, this tells us that the three planes given by the original system do not all intersect at any common points.

- (c) When $a = 1$ and $b = 12$ our augmented matrix reduces to

$$\left[\begin{array}{ccc|c} 1 & 1 & 0 & 2 \\ 0 & 0 & 1 & -\frac{1}{3} \\ 0 & 0 & 0 & 0 \end{array} \right].$$

There are no rows that correspond to equations of the form $0 = c$ for a nonzero constant c , so the system is consistent. The variable y is a free variable, so the system has infinitely many solutions. Algebraically, the solutions are y is free, is $z = -\frac{1}{3}$, and $x = 2 - y$. Geometrically, this tells us that the three planes given by the original system intersect in the line with $z = -\frac{1}{3}$, and $x = 2 - y$.



Summary

- A matrix is just a rectangular array of numbers or objects.
- Given a system of linear equations, with the variables listed in the same order in each equation, we represent the system by writing the coefficients of the first equation as the first row of a matrix, the coefficients of the second equation as the second row, and so on. This creates the coefficient matrix of the system. We then augment the coefficient matrix with a column of the constants that appear in the equations. This gives us the augmented matrix of the system.
- The operations that we can perform on equations translate exactly to row operations that we can perform on an augmented matrix:
 - (1) Replacing one row by the sum of that row and a scalar multiple of another row.
 - (2) Interchanging two rows.
 - (3) Replacing a row by a nonzero scalar multiple of itself.
- The forward elimination phase of the elimination method recursively eliminates the variables in a linear system to reach an equivalent but simplified system.
- The first non-zero entry in an equation in a linear system after elimination is called a pivot.
- A basic variable in a linear system corresponds to a pivot of the system. A free variable is a variable that is not basic.
- A linear system can be inconsistent (no solutions), have a unique solution (if consistent and every variable is a basic variable), or have infinitely many solutions (if consistent and there is a free variable).
- A linear system has no solutions if, after elimination, there is an equation of the form $0 = b$ where b is a nonzero number.
- A linear system after the elimination method can be solved using back-substitution. The free variables can be chosen arbitrarily and the basic variables can be solved in terms of the free variables through the back-substitution process.

Exercises

- (1) Consider the system of linear equations whose augmented matrix is

$$\left[\begin{array}{cc|c} 1 & 3 & -1 \\ 2 & h & k \end{array} \right]$$

where h and k are unknown constants. For which values of h and k does this system have

- a unique solution,
- infinitely many solutions,
- no solution?



- (2) Consider the following system:

$$\begin{aligned}x - 2y + z &= -1 \\-x + y - 3z &= 2 \\x + hy - z &= 0.\end{aligned}$$

Check that when $h = -3$ the system has infinitely many solutions, while when $h \neq -3$ the system has a unique solution.

- (3) If possible, find a system of three equations (not in reduced form) in three variables whose solution set consists only of the point $x_1 = 2, x_2 = -1, x_3 = 0$.
- (4) What are the possible geometrical descriptions of the solution set of two linear equations in \mathbb{R}^3 ? (Recall that \mathbb{R}^3 is the three-dimensional xyz -space – that is, the set of all ordered triples of the form (x, y, z)).
- (5) Two students are talking about when a linear system has infinitely many solutions.

Student 1: So, if we have a linear system whose augmented matrix has a row of zeros, then the system has infinitely many solutions, doesn't it?

Student 2: Well, but what if there is a row of the form $[0\ 0 \dots 0 \mid b]$ with a non-zero b right above the row of 0's?

Student 1: OK, maybe I should ask “If we have a consistent linear system whose augmented matrix has a row of zeros, then the system has infinitely many solutions, doesn't it?”

Student 2: I don't know. It still doesn't sound enough to me, but I'm not sure why.

Is Student 1 right? Or is Student 2's hunch correct? Justify your answer with a specific example if possible.

- (6) Label each of the following statements as True or False. Provide justification for your response.
- (a) **True/False** A system of linear equations in two unknowns can have exactly five solutions.
 - (b) **True/False** A system of equations with all the right hand sides equal to 0 has at least one solution.
 - (c) **True/False** A system of equations where there are fewer equations than the number of unknowns (known as an underdetermined system) cannot have a unique solution.
 - (d) **True/False** A system of equations where there are more equations than the number of unknowns (known as an overdetermined system) cannot have a unique solution.
 - (e) **True/False** A consistent system of two equations in three unknowns cannot have a unique solution.
 - (f) **True/False** If a system with three equations and three unknowns has a solution, then the solution is unique.



- (g) **True/False** If a system of equations has two different solutions, then it has infinitely many solutions.
- (h) **True/False** If there is a row of zeros in the row echelon form of the augmented matrix of a system of equations, the system has infinitely many solutions.
- (i) **True/False** If there is a row of zeros in the row echelon form of the augmented matrix of a system of n equations in n variables, the system has infinitely many solutions.
- (j) **True/False** If a system has no free variables, then the system has a unique solution.
- (k) **True/False** If a system has a free variable, then the system has infinitely many solutions.

Project: A Polynomial Fitting Application: Simpson's Rule

As discussed in the introduction, Simpson's Rule for approximating a definite integral models the integrand with a quadratic polynomial on each interval. To better understand this method, we consider how to fit a quadratic to three points.

Suppose we are given a collection of three points in the plane: (x_1, y_1) , (x_2, y_2) and (x_3, y_3) . There is exactly one quadratic polynomial $p(x)$ which goes through these points, i.e. there is exactly one quadratic $p(x)$ such that for each x_i , $p(x_i) = y_i$. This is an example of *polynomial curve fitting*.

Suppose our given points are $(-1, 2)$, $(1, 6)$, $(2, 5)$. To fit a quadratic to these points, consider a general quadratic of the form $p(x) = a_2x^2 + a_1x + a_0$. By substituting the x value of each of the given points and setting that equal to the y value of that point, we find three equations

$$(-1)^2a_2 - a_1 + a_0 = 2, \quad a_2 + a_1 + a_0 = 6, \quad (2)^2a_2 + 2a_1 + a_0 = 5$$

that give us a system of three equations in the three unknowns a_2 , a_1 , and a_0 :

$$\begin{aligned} a_2 - a_1 + a_0 &= 2 \\ a_2 + a_1 + a_0 &= 6 \\ 4a_2 + 2a_1 + a_0 &= 5. \end{aligned}$$

This system is the example we considered in Preview Activity 2.1, whose solution is $a_2 = -1$, $a_1 = 2$, and $a_0 = 5$. A graph of $q(x) = -x^2 + 2x + 5$ along with the three points $(-1, 2)$, $(1, 6)$, $(2, 5)$ is shown in Figure 2.1.

Project Activity 2.1. In this activity we model the sine function on the interval $[a, b]$, where $a = -\frac{\pi}{2}$ and $b = \pi$ with a collection of quadratics. Let $f(x) = \sin(x)$. We partition the interval $[a, b]$ using 6 partition points. Let $x_0 = -\frac{\pi}{2}$, $x_1 = -\frac{\pi}{4}$, $x_2 = 0$, $x_3 = \frac{\pi}{4}$, $x_4 = \frac{\pi}{2}$, $x_5 = \frac{3\pi}{4}$, and $x_6 = \pi$. We need 3 points to determine a quadratic, so the interval $[a, b]$ will be partitioned into 3 subintervals: $[x_0, x_2]$, $[x_2, x_4]$, and $[x_4, x_6]$.

- (a) Set up a system of linear equations to fit a quadratic $q_1(x) = r_1x^2 + s_1x + t_1$ to the 3 points $(x_0, f(x_0))$, $(x_1, f(x_1))$, and $(x_2, f(x_2))$. (The solution to this system to 3 decimal places is $r_1 \approx 0.336$, $s_1 \approx 1.164$, and $t_1 = 0$.)



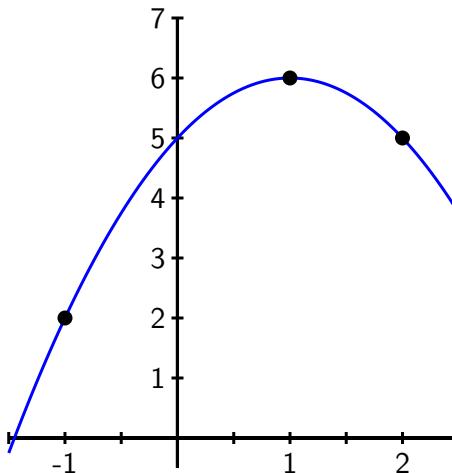


Figure 2.1: A quadratic fit to the points $(-1, 2)$, $(1, 6)$, $(2, 5)$.

- (b) Set up a system of linear equations to fit a quadratic $q_2(x) = r_2x^2 + s_2x + t_2$ to the 3 points $(x_2, f(x_2))$, $(x_3, f(x_3))$, and $(x_4, f(x_4))$. (The solution to this system to 3 decimal places is $r_2 \approx -0.336$, $s_2 \approx 1.164$, and $t_2 = 0$.)
- (c) Set up a system of linear equations to fit a quadratic $q_3(x) = r_3x^2 + s_3x + t_3$ to the 3 points $(x_4, f(x_4))$, $(x_5, f(x_5))$, and $(x_6, f(x_6))$. (The solution to this system to 3 decimal places is $r_3 \approx -0.336$, $s_3 \approx 0.946$, and $t_3 \approx 0.343$.)
- (d) Plot the three quadratics on their intervals against the graph of f . Explain what you see.

Project Activity 2.1 illustrates how we can model a function on an interval using a sequence of quadratic functions. Now we apply this polynomial curve fitting technique to derive the general formula for Simpson's Rule for approximating definite integrals. The Simpson sum $S(n)$ is found by using parabolic arcs to approximate the graph of f on each subinterval rather than line segments. This allows Simpson's Rule to more closely approximate the value of the definite integral with a smaller number of subintervals, although Simpson's Rule requires more calculations. Recall that to approximate a definite integral of a function f on an interval $[a, b]$, we partition $[a, b]$ into equal length subintervals. For Simpson's Rule, we partition $[a, b]$ into $n = 2m$ subintervals of equal length $\Delta x = \frac{b-a}{n}$. (Note that we need an even number of subintervals since we have to use three points for each parabola.) For each k we let $x_k = a + k\Delta x$ and $y_k = f(x_k)$. We approximate f on each subinterval using a quadratic. So we need to find the quadratic $Q(x) = c_2x^2 + c_1x + c_0$ that passes through two consecutive end points as well as the midpoint of a subinterval. That is, we need to find the coefficients of Q so that Q passes through the points (x_k, y_k) , (x_{k+2}, y_{k+2}) , and the midpoint (x_{k+1}, y_{k+1}) on the interval $[x_k, x_{k+2}]$ (so that we have three points to which to fit a parabola). Note that the length of the interval $[x_k, x_{k+2}]$ is $2\Delta x$. To make the calculations easier, we will translate our function so that our leftmost point is $(-r, y_k)$. Then the middle point is $(0, y_{k+1})$ and the rightmost point is (r, y_{k+2}) , where $r = \Delta x$.

Project Activity 2.2.

- (a) Set up a linear system that will determine the coefficients c_2 , c_1 , and c_0 so that the polyno-

mial $Q(x) = c_2x^2 + c_1x + c_0$ passes through the points $(-r, y_k)$, $(0, y_{k+1})$, and (r, y_{k+2}) with $r \neq 0$. Remember that the unknowns in this system are c_2 , c_1 , and c_0 .

- (b) We apply row operations to the matrix $\begin{bmatrix} r^2 & -r & 1 & y_k \\ 0 & 0 & 1 & y_{k+1} \\ r^2 & r & 1 & y_{k+2} \end{bmatrix}$ and obtain the matrix $\begin{bmatrix} r^2 & -r & 1 & y_k \\ 0 & 2r & 0 & y_{k+2} - y_k \\ 0 & 0 & 1 & y_{k+1} \end{bmatrix}$. Use these matrices to show that $c_2 = \frac{y_k - 2y_{k+1} + y_{k+2}}{2r^2}$, $c_1 = \frac{y_{k+2} - y_k}{2r}$, and $c_0 = y_{k+1}$.
- (c) Our goal is to ultimately approximate $\int_a^b f(x) dx$ by approximating f with quadratics on each subinterval. Use the fact that

$$\int_{x_k}^{x_{k+2}} f(x) dx \approx \int_{x_k}^{x_{k+2}} Q(x) dx = \int_{-r}^r Q(x) dx,$$

to show that

$$\int_{x_k}^{x_{k+2}} f(x) dx = \frac{1}{3} (y_k + 4y_{k+1} + y_{k+2}) \Delta x.$$

- (d) Now we can derive Simpson's Rule. Use an additive property of the definite integral to show that

$$\int_a^b f(x) dx \approx S(n),$$

where

$$S(n) = (y_0 + 4y_1 + 2y_2 + 4y_3 + 2y_4 + \cdots + 2y_{n-2} + 4y_{n-1} + y_n) \frac{\Delta x}{3}$$

is the Simpson's Rule approximation.

Notice that we can rewrite the Simpson's Rule approximation as

$$\begin{aligned} & \frac{1}{3} (y_0 + 4y_1 + 2y_2 + 4y_3 + 2y_4 + \cdots + 2y_{n-2} + 4y_{n-1} + y_n) \Delta x \\ &= \frac{2}{3} [2(y_1 + y_3 + \cdots + y_{n-1})] \Delta x \\ &\quad + \frac{2}{3} \left(\frac{y_0 + y_2}{2} + \frac{y_2 + y_4}{2} + \cdots + \frac{y_{n-2} + y_n}{2} \right) \Delta x \\ &= \frac{1}{3} [(y_1 + y_3 + \cdots + y_{n-1})(2\Delta x)] \\ &\quad + \frac{1}{3} \left(\frac{y_0 + y_2}{2} + \frac{y_2 + y_4}{2} + \cdots + \frac{y_{n-2} + y_n}{2} \right) (2\Delta x) \\ &= \frac{2M(n) + T(n)}{3}, \end{aligned}$$

where $M(n)$ is the midpoint sum and $T(n)$ is the trapezoid sum using n subdivisions of the interval $[a, b]$. Therefore, the weighted average $\frac{2M(n)+T(n)}{3}$ of the midpoint and trapezoid sums gives an approximation using quadratic functions.

Section 3

Row Echelon Forms

Focus Questions

By the end of this section, you should be able to give precise and thorough answers to the questions listed below. You may want to keep these questions in mind to focus your thoughts as you complete the section.

- What is the row echelon form of a matrix?
- What is the procedure to obtain the row echelon form of any matrix?
- What is the reduced row echelon form of a matrix?
- What is the procedure to obtain the reduced row echelon form of any matrix?
- What do the echelon forms of the augmented matrix for a linear system tell us about the solutions to the system?

Application: Balancing Chemical Reactions

Linear systems have applications in chemistry when balancing chemical equations. When a chemical reaction occurs, molecules of different substances combine to create molecules of other substances. Chemists represent such reactions with chemical equations. To balance a chemical equation means to find the number of atoms of each element involved that will preserve the number of atoms in the reaction. As an example, consider the chemical equation



This equation asks about what will happen when the chemicals ethane (C_2H_6) and oxygen (O_2), called the *reactants* of the reaction, combine to produce carbon dioxide (CO_2) and water (H_2O), called the *products* of the reaction (note that oxygen gas is *diatomic*, so that oxygen atoms are paired). The arrow indicates that it is the reactants that combine to form the products. Any chemical reaction has to obey the Law of Conservation of Mass that says that mass can neither be created

nor destroyed in a chemical reaction. Consequently, a chemical reaction requires the same number of atoms on both sides of the reaction. In other words, the total mass of the reactants must equal the total mass of the products. In reaction (3.1) the chemicals involved are made up of carbon (C), hydrogen (H), and oxygen (O) atoms. To balance the equation, we need to know how many molecules of each chemical are combined to preserve the number of atoms of C, H, and O. This can be done by setting up a linear system of equations of the form

$$\begin{aligned} 2x_1 - x_3 &= 0 \\ 6x_1 &- 2x_4 = 0 \\ 2x_2 - 2x_3 - x_4 &= 0, \end{aligned}$$

where x_1 , x_2 , x_3 , and x_4 represent the number of molecules of C_2H_6 , O_2 , CO_2 , and H_2O , respectively, in the reaction and then solving the system. Specific details can be found at the end of this section.

Introduction

In the previous sections, we identified operations on a given linear system with corresponding equivalent operations on the matrix representations which simplify the system and its matrix representation without changing the solutions of the system. Our end goal was to obtain a system which could be solved using back substitution, such as

$$\begin{aligned} x_1 - x_2 + x_3 &= 0 \\ 6x_2 - x_3 &= 8 \\ x_3 &= 1. \end{aligned}$$

The augmented matrix for this system is

$$\left[\begin{array}{ccc|c} 1 & -1 & 1 & 0 \\ 0 & 6 & -1 & 8 \\ 0 & 0 & 1 & 1 \end{array} \right].$$

The matrices of linear systems which can be solved via back substitution are said to be in *row echelon form* (or simply *echelon form*). We will define the properties of matrices in this form precisely in this section. Our goal will be to prescribe a precise procedure for converting any matrix to an equivalent one in row echelon form without having to convert back to the system representation.

Preview Activity 3.1. We want to determine a suitable form for an augmented matrix that can be obtained from row operations so that it is straightforward to find the solutions to the system. We begin with some examples.

- (1) Write the linear system corresponding to each of the following augmented matrices. Use the linear system to determine which systems have their variables eliminated completely in the forward direction, or equivalently determine for which systems the next step in the solution



process is back substitution (possibly using free variables). Explain your reasoning. You do not need to solve the systems.

$$\begin{array}{ll} \text{i. } \left[\begin{array}{ccc|c} 1 & -1 & 2 & -2 \\ 0 & 1 & 2 & -1 \\ 0 & 0 & 3 & 1 \end{array} \right] & \text{ii. } \left[\begin{array}{ccc|c} 1 & 1 & 0 & -2 \\ 0 & 1 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{array} \right] \\ \text{iii. } \left[\begin{array}{ccc|c} 1 & 1 & 1 & 2 \\ 1 & 2 & 2 & 2 \\ 0 & 0 & 2 & 2 \end{array} \right] & \text{iv. } \left[\begin{array}{ccc|c} 0 & 1 & 1 & 2 \\ 0 & 0 & 3 & 3 \\ 0 & 0 & -2 & -2 \end{array} \right] \end{array}$$

- (2) Shown below are two row reduced forms of the system

$$\begin{aligned} 2x_1 - x_3 &= 0 \\ 6x_1 &- 2x_4 = 0 \\ 2x_2 - 2x_3 - x_4 &= 0. \end{aligned}$$

Of the systems that correspond to these augmented matrices, which is easier to solve and why?

$$\left[\begin{array}{cccc|c} 2 & 0 & -1 & 0 & 0 \\ 0 & 2 & -2 & -1 & 0 \\ 0 & 0 & 3 & -2 & 0 \end{array} \right] \quad \left[\begin{array}{cccc|c} 1 & 0 & 0 & -\frac{1}{3} & 0 \\ 0 & 1 & 0 & -\frac{7}{6} & 0 \\ 0 & 0 & 1 & -\frac{2}{3} & 0 \end{array} \right]$$

The Echelon Forms of a Matrix

In the previous sections we saw how to simplify a linear system and its matrix representation via the elimination method without changing the solution set. This process is more efficient when performed on the matrix representation rather than on the system itself. Furthermore, the process of applying row operations to any augmented matrix is one that can be automated. In order to write an algorithm that can be used with any size augmented matrix to the extent that it can be applied even by a computer program, it is necessary to have a consistent procedure and a stopping point for the simplification process. The two main properties that we want the simplified augmented matrix to satisfy are that it should be easy to see if the system has solutions from the simplified matrix, and in cases when there are solutions, the general form of the solutions can be easily found. Hence the topic of this section is to define the process of elimination completely and generally.

We begin by discussing the *row echelon* or, simply, *echelon* form of a matrix. We know that the forward phase of the elimination on a linear system produces a system which can be solved by back substitution. The matrix representation of such a simplified system is said to be in *row echelon* or simply *echelon* form. Note that matrices in this form have the first nonzero entry in each row to the right of and below the first nonzero entry in the preceding row. Our next step is to formally describe this form – one that you tried to explain in problem 3 of Preview Activity 3.1.

Definition 3.1. A rectangular matrix is in **row echelon form** (or simply **echelon form**) if it has the following properties:

- (1) All nonzero rows are above any rows of all zeros.



- (2) Each **pivot** (the first non-zero entry reading from left to right) in a row is in a column to the right of the pivot of the row above it.

A pivot is also called a *leading entry* of a row. Note that properties (1) and (2) above imply that all entries in a column below a pivot are zeros. It can be shown that the positions of these pivots, called **pivot positions**, are unique and tell us quite a bit about a matrix and the solutions of the linear system it corresponds to. The columns that the pivots are in, called **pivot columns**, will also have useful properties as we will see soon.

Reflection 3. Compare the row echelon form of an augmented matrix to the corresponding system. Do you clearly see the correspondence between the requirements of the row echelon form and the properly eliminated variables in the system? Can you quickly come up with a system which will be in row echelon form when represented in augmented matrix form? Can you modify the standard row echelon form definition to cover cases where the elimination process eliminates the variables from last to first? For example, in a system with three equations in three unknowns, the last variable, say x_3 , can be eliminated from the second equation, and the last two variables, say x_2, x_3 can be eliminated from the last equation. How would you define this modified row echelon form for a general system with this modified elimination process?

Once an augmented matrix is in row echelon form, we can use back substitution to solve the corresponding system. However, we can make solving much easier with just a little more elimination work.

Row operations are easy to apply, so if we are inclined, there is no reason to stop at the row echelon form. For example, starting with the following matrix

$$\left[\begin{array}{cccc|c} 2 & -1 & 2 & 2 & 7 \\ 0 & 1 & 3 & -1 & -1 \\ 0 & 0 & 0 & 2 & 4 \end{array} \right]$$

in row echelon form, we could take the row operations even farther and avoid the process of back substitution altogether. First, we multiply the last row by $1/2$ to simplify that row:

$$\frac{1}{2}R_3 \rightarrow R_3 \quad \left[\begin{array}{cccc|c} 2 & -1 & 2 & 2 & 7 \\ 0 & 1 & 3 & -1 & -1 \\ 0 & 0 & 0 & 1 & 2 \end{array} \right].$$

Then we use the third row to eliminate entries above the third pivot:

$$\begin{aligned} R_1 - 2R_3 \rightarrow R_1 &\quad \left[\begin{array}{cccc|c} 2 & -1 & 2 & 0 & 3 \\ 0 & 1 & 3 & 0 & 1 \\ 0 & 0 & 0 & 1 & 2 \end{array} \right] \\ R_2 + R_3 \rightarrow R_2 &\quad \left[\begin{array}{cccc|c} 2 & -1 & 2 & 0 & 3 \\ 0 & 1 & 3 & 0 & 1 \\ 0 & 0 & 0 & 1 & 2 \end{array} \right]. \end{aligned}$$

We can continue in this manner (we call this process *backward elimination*) to make 0 all of the entries above the pivots (one in the second column, and one in the fourth) with the pivots being 1, to ultimately obtain the equivalent augmented matrix

$$\left[\begin{array}{cccc|c} 1 & 0 & 1 & 0 & 2 \\ 0 & 1 & 3 & 0 & 1 \\ 0 & 0 & 0 & 1 & 2 \end{array} \right].$$



The system corresponding to this augmented matrix is

$$\begin{array}{rcl} x_1 + x_3 & = & 2 \\ x_2 + 3x_3 & = & 1 \\ x_4 & = & 2 \end{array}$$

so we can just directly read off the solution to the system: x_3 free and $x_1 = 2 - x_3, x_2 = 1 - 3x_3, x_4 = 2$. This final row reduced form makes solving the system very easy, and this form is called the *reduced row echelon form*.

Definition 3.2. A rectangular matrix is in **reduced row echelon form** (or **reduced echelon form**) if the matrix is in row echelon form and

- (3) The pivot in each nonzero row is 1.
- (4) Each pivot is the only nonzero entry in its column.

In short, the reduced row echelon form of a matrix is a row echelon form in which all the pivots are 1 and any entries below and above the pivots are 0.

If we use either of these two row echelon forms, solving the original system becomes straightforward and, as a result, these matrix forms are stopping points for the row operation algorithm to solve a system. It is also very easy to write a computer program to perform row operations to obtain a row echelon or reduced row echelon form of the matrix, making hand computations unnecessary. We will discuss this shortly.

Reflection 4. Compare the reduced row echelon form of an augmented matrix to the corresponding system. Do you clearly see the correspondence between the requirements of the reduced row echelon form and the way the variables appear in the equations in the system? Can you quickly come up with a system which will be in reduced row echelon form when represented in augmented matrix form?

Note. We have used the elimination method on augmented matrices so far. However, the elimination method can be applied on just the coefficient matrix, or other matrices that will arise in other contexts, and will provide useful information in each of those cases. Therefore, the row echelon form and reduced row echelon form is defined for any matrix, and from now on, a matrix will be a general matrix unless explicitly specified to be an augmented matrix.

Activity 3.1. Identify which of the following matrices is in row echelon form (REF) and/or reduced row echelon form (RREF). For those in row and/or reduced row echelon form, identify the pivots clearly by circling them. For those that are not in a given form, state which properties the matrix fails to satisfy.

(a) $\left[\begin{array}{cccc} 2 & 4 & -3 & 6 \\ 0 & 0 & 0 & 7 \end{array} \right]$ (b) $\left[\begin{array}{cc} 1 & 0 \\ 0 & 1 \end{array} \right]$ (c) $\left[\begin{array}{cccc} 0 & 1 & 2 & 3 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 5 \end{array} \right]$

(d) $\left[\begin{array}{cccc} 1 & 2 & 3 & 4 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array} \right]$ (e) $\left[\begin{array}{cc} 0 & 0 \\ 0 & 0 \end{array} \right]$

Determining the Number of Solutions of a Linear System

Consider the system

$$\begin{array}{rcl} x_1 + 2x_2 - x_3 & = & 0 \\ x_2 & - & x_4 = 2 \\ x_3 - 2x_4 & = & 4. \end{array}$$

The augmented matrix for this system is

$$\left[\begin{array}{cccc|c} 1 & 2 & -1 & 0 & 0 \\ 0 & 1 & 0 & -1 & 2 \\ 0 & 0 & 1 & -2 & 4 \end{array} \right].$$

Note that this matrix is already in row echelon form. The reduced row echelon form of this augmented matrix is

$$\left[\begin{array}{cccc|c} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & -1 & 2 \\ 0 & 0 & 1 & -2 & 4 \end{array} \right]. \quad (3.2)$$

Since there are leading 1s in the first three columns, we can use those entries to write x_1 , x_2 , and x_3 in terms of x_4 . We then choose x_4 to be arbitrary and write the remaining variables in terms of x_4 . Let $x_4 = t$. Solving the third equation for x_3 gives us $x_3 = 4 + 2t$. The second equation shows that $x_2 = 2 + t$, and the first that $x_1 = 0$. Each value of t provides a solution to the system, so our system has infinitely many solutions. These solutions are

$$x_1 = 0, x_2 = 2 + t, x_3 = 4 + 2t, \text{ and } x_4 = t,$$

where t can have any value.

Activity 3.2. We have seen examples of systems with no solutions, one solution, and infinitely many solutions. As we will see in this activity, we can recognize the number of solutions to a system by analyzing the pivot positions in the augmented matrix of the system.

- (a) Write an example of an augmented matrix in row echelon form so that the last column of the (whole) matrix is a pivot column. What is the system of equations corresponding to your augmented matrix? How many solutions does your system have? Why?
- (b) Consider the reduced row echelon form (3.2). Based on the columns of this matrix, explain how we know that the system it represents is consistent.
- (c) The system with reduced row echelon form (3.2) is consistent. What is it about the columns of the coefficient matrix that tells us that this system has infinitely many solutions?
- (d) Suppose that a linear system is consistent and that the coefficient matrix has m rows and n columns.
 - i. If every column of the coefficient matrix is a pivot column, how many solutions must the system have? Why? What relationship must exist between m and n ? Explain.



- ii. If the coefficient matrix has at least one non-pivot column, how many solutions must the system have? Why?

When solving a linear system of equations, the free variables can be chosen arbitrarily and we can write the basic variables in terms of the free variables. Therefore, the existence of a free variable leads to infinitely many solutions for consistent systems. However, it is possible to have a system with free variables which is inconsistent. (Can you think of an example?)

Producing the Echelon Forms

In this part, we consider the formal process of creating the row and reduced row echelon forms of matrices. The process of creating the row echelon form is the equivalent of the elimination method on systems of linear equations.

Activity 3.3. Each of the following matrices is at most a few steps away from being in the requested echelon form. Determine what row operations need to be completed to turn the matrix into the required form.

(a) Turn into REF: $\begin{bmatrix} 0 & 2 \\ 2 & 1 \end{bmatrix}$	(b) Turn into REF: $\begin{bmatrix} 1 & 2 \\ 2 & 5 \end{bmatrix}$
(c) Turn into RREF: $\begin{bmatrix} 2 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 1 \end{bmatrix}$	(d) Turn into RREF: $\begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix}$
(e) Turn into RREF: $\begin{bmatrix} 1 & 1 \\ 0 & 2 \end{bmatrix}$	(f) Turn into RREF: $\begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 3 \\ 0 & 0 & 2 \end{bmatrix}$

The complete process of applying row operations to reduce an augmented matrix to a row or reduced row echelon form can be expressed as a recursive process in an algorithmic fashion, making it possible to program computers to solve linear systems. Here are the steps to do so:

Step 1: Begin with the leftmost nonzero column (if there is one). This will be a pivot column.

Step 2: Select a nonzero entry in this pivot column as a pivot. If necessary, interchange rows to move this entry to the first row (this entry will be a pivot).

Step 3: Use row operations to create zeros in all positions below the pivot.

Step 4: Cover (or ignore) the row containing the pivot position and cover all rows, if any, above it. Apply steps 1-3 to the submatrix that remains. Repeat the process until there are no more nonzero rows to modify.

To obtain the reduced row echelon form we need one more step.

Step 5: Beginning with the rightmost pivot and working upward and to the left, create zeros above each pivot. If a pivot is not 1, make it 1 by an appropriate row multiplication.



The algorithm described in steps 1-4 will produce the row echelon form of the matrix. This algorithm is called *Gaussian elimination*. When we add step 5 to produce the reduced row echelon form, the algorithm is called *Gauss-Jordan elimination*.

Activity 3.4. Consider the matrix $\begin{bmatrix} 0 & 2 & 4 & 1 \\ -1 & 3 & 0 & 6 \\ 0 & 4 & 8 & 2 \\ 1 & -3 & 0 & -2 \end{bmatrix}$.

- Perform Gaussian elimination to reduce the matrix to row echelon form. Clearly identify each step used. Compare your row echelon form to that of another group. Do your results agree? If not, who is right?
- Now continue applying row operations to obtain the reduced row echelon form of the matrix. Clearly identify each step. Compare your row echelon form to that of another group. Do your results agree? If not, who is right?

If we compare row echelon forms from Activity 3.4, it is likely that different groups or individuals produced different row echelon forms. That is because the row echelon form of a matrix is not unique. (Is the row echelon form ever unique?)

However, if row operations are applied correctly, then we will all arrive at the same reduced row echelon form in Activity 3.4:

$$\begin{bmatrix} 1 & 0 & 6 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

It turns out that the reduced row echelon form of a matrix is unique.

Two matrices who are connected by row operations are said to be *row equivalent*.

Definition 3.3. A matrix B is **row equivalent** to a matrix A if B can be obtained by applying elementary row operations to A .

Since every elementary row operation is reversible, if B is row equivalent to A , then A is also row equivalent to B . Thus, we just say that A and B are row equivalent. While the row echelon form of a matrix is not unique, it is the case that the reduced row echelon form of a matrix is unique.

Theorem 3.4. *Every matrix is row equivalent to a unique matrix in reduced row echelon form.*

The reduced row echelon form of a matrix that corresponds to a system of linear equations provides us with a equivalent system whose solutions are easy to find. As an example, consider the system

$$\begin{aligned} 2x_2 + 4x_3 + x_4 &= 0 \\ -x_1 + 3x_2 + 6x_4 &= 0 \\ 4x_2 + 8x_3 + 2x_4 &= 0 \\ x_1 - 3x_2 - 2x_4 &= 0 \end{aligned}$$



with augmented matrix

$$\left[\begin{array}{cccc|c} 0 & 2 & 4 & 1 & 0 \\ -1 & 3 & 0 & 6 & 0 \\ 0 & 4 & 8 & 2 & 0 \\ 1 & -3 & 0 & -2 & 0 \end{array} \right].$$

Notice that the coefficient matrix (the left hand side portion of the augmented matrix) of this system is same as the matrix we considered in Activity 3.4. Since we are augmenting with a column of zeros, no row operations will change those zeros in the augmented column. So the row operations applied in Activity 3.4 will give us the reduced row echelon form of this augmented matrix as

$$\left[\begin{array}{cccc|c} 1 & 0 & 6 & 0 & 0 \\ 0 & 1 & 2 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array} \right].$$

Note that the third column is not a pivot column. That means that the variable x_3 is a free variable. There are pivots in the other three columns of the coefficient matrix, so we can solve for x_1 , x_2 , and x_4 in terms of x_3 . These variables are the basic variables. The third row of the augmented matrix tells us that $x_4 = 0$. The second row corresponds to the equation $x_2 + 2x_3 = 0$, and solving for x_2 shows that $x_2 = -2x_3$. Finally, the first row tells us that $x_1 + 6x_3 = 0$, so $x_1 = -6x_3$. Therefore, the general solution to this system of equations is

$$x_1 = -6x_3, \quad x_2 = -2x_3, \quad x_3 \text{ is free}, \quad x_4 = 0.$$

The fact that x_3 is free means that we can choose any value for x_3 that we like and obtain a specific solution to the system. For example, if $x_3 = -1$, then we have the solution $x_1 = 6$, $x_2 = 2$, $x_3 = -1$, and $x_4 = 0$. Check this to be sure.

Activity 3.5. Each matrix below is an augmented matrix for a linear system after elimination with variables x_1, x_2, \dots in that order. Identify the basic variables (if any) and free variables (if any). Then find the general solution (if there is a solution) expressing all variables in terms of the free variables.

(a) $\left[\begin{array}{ccc c} 1 & 0 & 2 & 1 \\ 0 & 3 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{array} \right]$	(b) $\left[\begin{array}{ccc c} 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 \end{array} \right]$	(c) $\left[\begin{array}{cccc c} 1 & 2 & -1 & 1 & 1 \\ 0 & 1 & 0 & 2 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array} \right]$
(d) $\left[\begin{array}{ccc c} 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 2 \end{array} \right]$	(e) $\left[\begin{array}{cc c} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{array} \right]$	

Recall that in the previous section, we determined the criteria for when a system has a unique solution, or infinitely many solutions, or no solution. With the use of the row echelon form of the augmented matrix, we can rewrite these criteria as follows:

Theorem 3.5.

- (1) A linear system is consistent if in the row echelon form of the augmented matrix representing the system no pivot is in the rightmost column.



- (2) If a linear system is consistent and the row echelon form of the coefficient matrix does not have a pivot in every column, then the system has infinitely many solutions.
- (3) If a linear system is consistent and there is a pivot in every column of the row echelon form of the coefficient matrix, then the system has a unique solution.

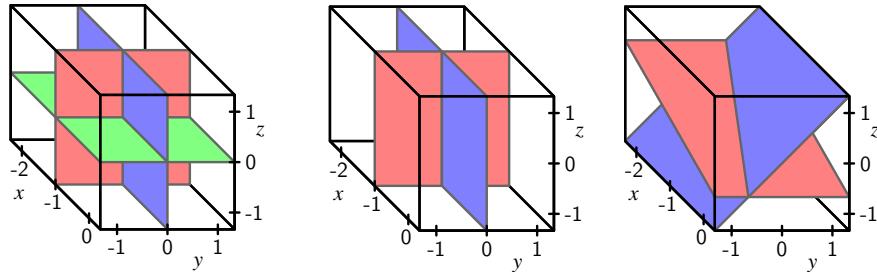


Figure 3.1: Figures for Activity 3.6.

Activity 3.6.

- (a) For each part, the reduced row echelon form of the augmented matrix of a system of equations in variables x , y , and z (in that order) is given. Use the reduced row echelon form to find the solution set to the original system of equations.

$$\text{i. } \left[\begin{array}{ccc|c} 1 & 0 & 0 & -1 \\ 0 & 1 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{array} \right] \text{ ii. } \left[\begin{array}{ccc|c} 1 & 0 & 2 & -1 \\ 0 & 1 & -1 & 3 \\ 0 & 0 & 0 & 0 \end{array} \right] \text{ iii. } \left[\begin{array}{ccc|c} 1 & 0 & 0 & 2 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & 3 \end{array} \right]$$

- iv. Each of the three systems above is represented as one of the graphs in Figure 3.1. Match each figure with a system.

- (b) The reduced row echelon form of the augmented matrix of a system of equations in variables x , y , z , and t (in that order) is given. Use the reduced row echelon form to find the solution set to the original system of equations:

$$\left[\begin{array}{cccc|c} 1 & 3 & 0 & 0 & -1 \\ 0 & 0 & 1 & 2 & 4 \\ 0 & 0 & 0 & 0 & 1 \end{array} \right].$$

Examples

What follows are worked examples that use the concepts from this section.

Example 3.6. Consider the linear system

$$\begin{aligned} 2x_1 + 6x_3 &= x_2 + 2 \\ 2x_3 - 4x_1 &= 2x_2 \\ x_2 + 4x_3 - 2 &= 2x_1 + 6. \end{aligned}$$

- (a) Find the augmented matrix for this system.
- (b) Use row operations to find a row echelon form of the augmented matrix of this system.
- (c) Use row operations to find the reduced row echelon form of the augmented matrix of this system.
- (d) Find the solution(s), if any, to the system.

Example Solution. Before we can find the augmented matrix of this system, we need to rewrite the system so that the variables are all on one side and the constant terms are on the other side of the equations. Doing so yields the equivalent system

$$\begin{aligned} 2x_1 - x_2 + 6x_3 &= 2 \\ -4x_1 - 2x_2 + 2x_3 &= 0 \\ -2x_1 + x_2 + 4x_3 &= 8. \end{aligned}$$

Note that this is not the only way to rearrange the system. For example, for the second equation, could be written instead as $4x_1 + 2x_2 - 2x_3 = 0$ to minimize the number of negative signs in the equation.

- (a) The augmented matrix for this system is

$$\left[\begin{array}{ccc|c} 2 & -1 & 6 & 2 \\ -4 & -2 & 2 & 0 \\ -2 & 1 & 4 & 8 \end{array} \right].$$

- (b) Our first steps to row echelon form are to eliminate the entries below the leading entry in the first row. To do this we replace row two with row two plus 2 times row 1 and we replace row three with row three plus row one. This produces the row equivalent matrix

$$\left[\begin{array}{ccc|c} 2 & -1 & 6 & 2 \\ 0 & -4 & 14 & 4 \\ 0 & 0 & 10 & 10 \end{array} \right].$$

This matrix is now in row echelon form.

- (c) To continue to find the reduced row echelon form, we replace row two with row two times $-\frac{1}{4}$ to get a leading 1 in the second row, and we replace row three with row three times $\frac{1}{10}$ to get a leading 1 in the third row and obtain the row equivalent matrix

$$\left[\begin{array}{ccc|c} 2 & -1 & 6 & 2 \\ 0 & 1 & -\frac{7}{2} & -1 \\ 0 & 0 & 1 & 1 \end{array} \right].$$

Now we perform backwards elimination to make the entries above the leading 1s equal to 0, starting with the third column and working backwards. Replace row one with row one

minus 6 times row three and replace row two with row two plus $\frac{7}{2}$ row three to obtain the row equivalent matrix

$$\left[\begin{array}{ccc|c} 2 & -1 & 0 & -4 \\ 0 & 1 & 0 & \frac{5}{2} \\ 0 & 0 & 1 & 1 \end{array} \right].$$

For the second column, we replace row one with row one plus row two to obtain the row equivalent matrix

$$\left[\begin{array}{ccc|c} 2 & 0 & 0 & -\frac{3}{2} \\ 0 & 1 & 0 & \frac{5}{2} \\ 0 & 0 & 1 & 1 \end{array} \right].$$

Since the leading entry in row one is not a one, we have one more step before we have the reduced row echelon form. Finally, we replace row one with row one times $\frac{1}{2}$. This gives us the reduced row echelon form

$$\left[\begin{array}{ccc|c} 1 & 0 & 0 & -\frac{3}{4} \\ 0 & 1 & 0 & \frac{5}{2} \\ 0 & 0 & 1 & 1 \end{array} \right].$$

- (d) We can read off the solution to the system from the reduced row echelon form: $x_1 = -\frac{3}{4}$, $x_2 = \frac{5}{2}$, and $x_3 = 1$. You should check in the original equations to make sure we have the correct solution.

Example 3.7. In this example, a and b are unknown scalars. Consider the system with augmented matrix

$$\left[\begin{array}{ccc|c} 1 & 2 & a & 3 \\ 1 & 0 & 0 & b \\ 0 & 1 & 1 & 0 \end{array} \right].$$

Find all values of a and b so that the system has:

- (a) Exactly one solution (and find the solution)
- (b) No solutions
- (c) Infinitely many solutions (and find all solutions)

Example Solution. Let x_1 , x_2 , and x_3 be the variables corresponding to the first, second, and third columns, respectively, of the augmented matrix. To answer these questions, we row reduce the augmented matrix. We interchange rows one and two and then also rows two and three to obtain the matrix

$$\left[\begin{array}{ccc|c} 1 & 0 & 0 & b \\ 0 & 1 & 1 & 0 \\ 1 & 2 & a & 3 \end{array} \right].$$



Now we replace row three with row three minus row one to produce the row equivalent matrix

$$\left[\begin{array}{ccc|c} 1 & 0 & 0 & b \\ 0 & 1 & 1 & 0 \\ 0 & 2 & a & 3-b \end{array} \right].$$

Next, replace row three with row three minus 2 times row two. This yields the row equivalent matrix

$$\left[\begin{array}{ccc|c} 1 & 0 & 0 & b \\ 0 & 1 & 1 & 0 \\ 0 & 0 & a-2 & 3-b \end{array} \right].$$

We now have a row echelon form.

- (a) The system will have exactly one solution when the last row has the form $[0 \ 0 \ u \ v]$ where u is not zero. Thus, the system has exactly one solution when $a - 2 \neq 0$, or when $a \neq 2$. In this case, the solution is

$$\begin{aligned} x_3 &= \frac{3-b}{a-2}, \\ x_2 &= -x_3 = \frac{b-2}{a-2} \\ x_1 &= b. \end{aligned}$$

You should check to ensure that this solution is correct. The other cases occur when $a = 2$.

- (b) When $a = 2$ and $3 - b \neq 0$ (or $b \neq 3$), then we have a row of the form $[0 \ 0 \ 0 \ t]$, where t is not 0. In these cases there are no solutions.
- (c) When $a = 2$ and $b = 3$, then the last row is a row of all zeros. In this case, the system is consistent and x_3 is a free variable, so the system has infinitely many solutions. The solutions are

$$\begin{aligned} x_1 &= b \\ x_2 &= -x_3 \\ x_3 &\text{ is free.} \end{aligned}$$

You should check to ensure that this solution is correct.

Summary

In this section we learned about the row echelon and reduced row echelon forms of a matrix and some of the things these forms tell us about solutions to systems of linear equations.

- A matrix is in row echelon form if
 - (1) All nonzero rows are above any rows of all zeros.
 - (2) Each **pivot** (the first nonzero entry) of a row is in a column to the right of the pivot of the row above it.



- Once an augmented matrix is in row echelon form, we can use back substitution to solve the corresponding linear system.
- To reduce a matrix to row echelon form we do the following:
 - Begin with the leftmost nonzero column (if there is one). This will be a pivot column.
 - Select a nonzero entry in this pivot column as a pivot. If necessary, interchange rows to move this entry to the first row (this entry will be a pivot).
 - Use row operations to create zeros in all positions below the pivot.
 - Cover (or ignore) the row containing the pivot position and cover all rows, if any, above it. Apply the preceding steps to the submatrix that remains. Repeat the process until there are no more nonzero rows to modify.
- A matrix is in reduced row echelon form if it is in row echelon form and
 - (3) The pivot in each nonzero row is 1.
 - (4) Each pivot is the only nonzero entry in its column.
- To obtain the reduced row echelon form from the row echelon form, beginning with the rightmost pivot and working upward and to the left, create zeros above each pivot. If a pivot is not 1, make it 1 by an appropriate row multiplication.
- Both row echelon forms of an augmented matrix tell us about the number of solutions to the corresponding linear system.
 - A linear system is inconsistent if and only if a row echelon form of the augmented matrix of the system contains a row of the form

$$[0 \ 0 \ 0 \ \cdots \ 0 \ *],$$

where * is not zero. Another way to say this is that a linear system is inconsistent if and only if the last column of the augmented matrix of the system is a pivot column.

- A consistent linear system will have a unique solution if and only if each column but the last in the augmented matrix of the system is a pivot column. This is equivalent to saying that a consistent linear system will have a unique solution if and only if the consistent system has no free variables.
- A consistent linear system will have infinitely many solutions if and only if the coefficient matrix of the system contains a non-pivot column. In that case, the free variables corresponding to the non-pivot columns can be chosen arbitrarily and the basic variables corresponding to pivot columns can be written in terms of the free variables.
- A linear system can have no solutions, exactly one solution, or infinitely many solutions.



Exercises

- (1) Represent the following linear system in variables x_1, x_2, x_3 in augmented matrix form and use row reduction to find the general solution of the system.

$$\begin{aligned}x_1 + x_2 - x_3 &= 4 \\x_1 + 2x_2 + 2x_3 &= 3 \\2x_1 + 3x_2 - 3x_3 &= 11.\end{aligned}$$

- (2) Represent the following linear system in variables x_1, x_2, x_3 in augmented matrix form after rearranging the terms and use row reduction to find all solutions to the system.

$$\begin{aligned}x_1 - x_3 - 2x_2 &= 3 \\2x_3 + 2 &= x_1 + x_2 \\4x_2 + 2x_1 - 2 &= 5x_3.\end{aligned}$$

- (3) Check that the reduced row echelon form of the matrix

$$\left[\begin{array}{cccc} 1 & -1 & 3 & 2 \\ -1 & 2 & -4 & -1 \\ 2 & 0 & 6 & 8 \end{array} \right]$$

is

$$\left[\begin{array}{cccc} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & 1 \end{array} \right].$$

- (4) Consider the following system:

$$\begin{aligned}x - 2y + z &= -1 \\2y - 4z &= 6 \\hy - 2z &= 1.\end{aligned}$$

- (a) Find a row echelon form of the augmented matrix for this system.

- (b) For which values of h , if any, does the system have (i.) no solutions, (ii.) exactly one solution, (iii.) infinitely many solutions? Find the solutions in each case.

- (5) Find the general solution of the linear system corresponding to the following augmented matrix:

$$\left[\begin{array}{cccc|c} 1 & -1 & 2 & 1 & 2 \\ -1 & 2 & 2 & -1 & -5 \\ 1 & 1 & 10 & 2 & -1 \end{array} \right].$$

- (6) What are the conditions, if any, on the a, b, c values so that the following augmented matrix corresponds to a consistent linear system? How many solutions will the consistent system have? Explain.

$$\left[\begin{array}{ccc|c} 1 & 2 & 3 & a \\ 2 & 3 & 7 & b \\ -1 & -4 & -1 & c \end{array} \right].$$

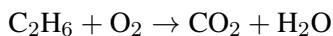
- (7) In this exercise the symbol \blacksquare denotes a non-zero number and the symbol * denotes any real number (including 0).
- Is the augmented matrix
- $$\left[\begin{array}{cc|c} \blacksquare & * & * \\ 0 & \blacksquare & * \end{array} \right]$$
- in a form to which back substitution will easily give the solutions to the system? Explain your reasoning. (Hint: In order to help see what happens in the general case, substitute some numbers in place of the \blacksquare 's and *'s and answer the question for that specific system first. Then determine if your answer generalizes.)
- The above matrix is a possible form of an augmented matrix with 2 rows and 3 columns corresponding to a linear system after forward elimination, i.e., a linear system for which back substitution will easily give the solutions. Determine the other possible such forms of the nonzero augmented matrices with 2 rows and 3 columns. As in part (a), use the symbol \blacksquare to denote a non-zero number and * to denote any real number.
 - Give an example of a linear system with a unique solution for which a row echelon form of the augmented matrix of the system has a row of 0's.
 - Come up with an example of an augmented matrix with 0's in the rightmost column corresponding to an inconsistent system, if possible. If not, explain why not.
 - Find two different row echelon forms which are equivalent to the same matrix not given in row echelon form.
 - Determine all possible row echelon forms of a 2×2 matrix. Use the symbol \blacksquare to denote a non-zero number and * to denote a real number with no condition on being 0 or not to represent entries.
 - Label each of the following statements as True or False. Provide justification for your response.
 - True/False** The number of pivots of an $m \times n$ matrix cannot exceed m . (Note: Here m, n are some unknown numbers.)
 - True/False** The row echelon form of a matrix is unique.
 - True/False** The reduced row echelon form of a matrix is unique.
 - True/False** A system of equations where there are fewer equations than the number of unknowns (known as an underdetermined system) cannot have a unique solution.
 - True/False** A system of equations where there are more equations than the number of unknowns (known as an overdetermined system) cannot have a unique solution.
 - True/False** If a row echelon form of the *augmented matrix* of a system of three equations in two unknowns has three pivots, then the system is inconsistent.
 - True/False** If the coefficient matrix of a system has pivots in every row, then the system is consistent.



- (h) **True/False** If there is a row of zeros in a row echelon form of the augmented matrix of a system of equations, the system has infinitely many solutions.
- (i) **True/False** If there is a row of zeros in a row echelon form of the augmented matrix of a system of n equations in n variables, the system has infinitely many solutions.
- (j) **True/False** If a linear system has no free variables, then the system has a unique solution.
- (k) **True/False** If a linear system has a free variable, then the system has infinitely many solutions.

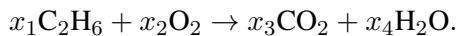
Project: Modeling a Chemical Reaction

Recall the chemical equation



from the beginning of this section. This equation illustrates the reaction between ethane (C_2H_6) and oxygen (O_2), called the *reactants*, to produce carbon dioxide (CO_2) and water (H_2O), called the *products* of the reaction. In any chemical reaction, the total mass of the reactants must equal the total mass of the products. In our reaction the chemicals involved are made up of carbon (C), hydrogen (H), and oxygen (O) atoms. To balance the equation, we need to know how many molecules of each chemical are combined to preserve the number of atoms of C, H, and O.

Let x_1 be the number of molecules of C_2H_6 , x_2 the number of molecules of O_2 , x_3 the number of molecules of CO_2 , and x_4 the number of molecules of H_2O in the reaction. We can then represent this reaction as



In each molecule (e.g., ethane C_2H_6), the subscripts indicate the number of atoms of each element in the molecule. So 1 molecule of ethane contains 2 atoms of carbon and 6 atoms of hydrogen. Thus, there are 2 atoms of carbon in C_2H_6 and 0 atoms of carbon in O_2 , giving us $2x_1$ carbon atoms in x_1 molecules of C_2H_6 and 0 carbon atoms in x_2 molecules of O_2 . On the product side of the reaction there is 1 carbon atom in CO_2 and 0 carbon atoms in H_2O . To balance the reaction, we know that the number of carbon atoms in the products must equal the number of carbon atoms in the reactants.

Project Activity 3.1.

- (a) Set up an equation that balances the number of carbon atoms on both sides of the reaction.
- (b) Balance the numbers of hydrogen and oxygen atoms in the reaction to explain why

$$\begin{aligned} 6x_1 &= 2x_4 \\ 2x_2 &= 2x_3 + x_4. \end{aligned}$$

- (c) So the system of linear equations that models this chemical reaction is

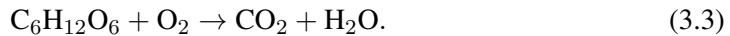
$$\begin{aligned} 2x_1 - x_3 &= 0 \\ 6x_1 - 2x_4 &= 0 \\ 2x_2 - 2x_3 - x_4 &= 0. \end{aligned}$$



Find all solutions to this system and then balance the reaction. Note that we cannot have a fraction of a molecule in our reaction. (Hint: Some of the work needed is done in Preview Activity 3.1.)

Project Activity 3.2. Chemical reactions can be very interesting.

- (a) Carbon dioxide, CO_2 , is a familiar product of combustion. For example, when we burn glucose, $\text{C}_6\text{H}_{12}\text{O}_6$, the products of the reaction are carbon dioxide and water:



Use the techniques developed in this project to balance this reaction.

- (b) To burn glucose, we need to add oxygen to make the combustion happen. Carbon dioxide is different in that it can burn without the presence of oxygen. For example, when we mix magnesium (Mg) with dry ice (CO_2), the products are magnesium oxide (MgO) and carbon (C). This is an interesting reaction to watch: you can see it at many websites, e.g., <http://www.ebaumsworld.com/video/watch/404311/> or <https://www.youtube.com/watch?v=-6dfi8LyRLA>.

Use the method determined above to balance the chemical reaction



Section 4

Vector Representation

Focus Questions

By the end of this section, you should be able to give precise and thorough answers to the questions listed below. You may want to keep these questions in mind to focus your thoughts as you complete the section.

- What is a vector?
- How do we define operations on vectors?
- What is a linear combination of vectors?
- How do we determine if one vector is a linear combination of a given set of vectors?
- How do we represent a linear system as a vector equation?
- What is the span of a set of vectors?
- What are possible geometric representations of the span of a vector, or the span of two vectors?

Application: The Knight's Tour

Chess is a game played on an 8×8 grid which utilizes a variety of different pieces. One piece, the knight, is different from the other pieces in that it can jump over other pieces. However, the knight is limited in how far it can move in a given turn. For these reasons, the knight is a powerful, but often under-utilized, piece.

A knight can move two units either horizontally or vertically, and one unit perpendicular to that. Four knight moves are as illustrated in Figure 4.1, and the other four moves are the opposites of these.

The knight's tour problem is the mathematical problem of finding a knight's tour, that is a se-

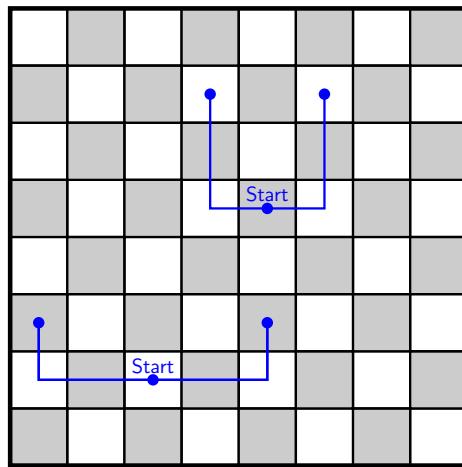


Figure 4.1: Moves a knight can make.

quence of knight moves so the the knight visits each square exactly once. While we won't consider a knight's tour in this text, we will see using linear combinations of vectors that a knight can move from its initial position to any other position on the board, and that it is possible to determine an sequence of moves to make that happen.

Introduction

So far we learned of a convenient method to represent a linear system using matrices. We now consider another representation of a linear system using *vectors*. Vectors can represent concepts in the physical world like velocity, acceleration, and force – but we will be interested in vectors as algebraic objects in this class. Vectors will form the foundation for everything we will do in linear algebra. For now, the following definition will suffice.

Definition 4.1. A (real) **vector** is a finite list of real numbers in a specified order. Each number in the list is referred to as an **entry** or **component** of the vector.

Note: For the majority of this text, we will work with real vectors. However, A vector does not need to be restricted to have real entries. At times we will use complex vectors and even vectors in other types of sets. The types of sets we use will be ones that have structure just like the real numbers. Recall that a real number is a number that has a decimal representation, either finite or repeating (rational numbers) or otherwise (irrational numbers). We can add and multiply real numbers as we have done throughout our mathematical careers, and the real numbers have a certain structure given in the following theorem that we will treat as an axiom – that is, we assume these properties without proof. We denote the set of real numbers with the symbol \mathbb{R} .

Theorem 4.2. Let x , y , and z be real numbers. Then

- $x + y \in \mathbb{R}$ and $xy \in \mathbb{R}$ (The name given to this property is closure. That is, the set \mathbb{R} is closed under addition and multiplication.)

- $x + y = y + x$ and $xy = yx$ (The name given to this property is commutativity. That is addition and multiplication are commutative operations in \mathbb{R} .)
- $(x+y)+z = x+(y+z)$ and $(xy)z = x(yz)$ (The name given to this property is associativity. That is, addition and multiplication is associative operations in \mathbb{R} .)
- There is an element 0 in \mathbb{R} such that $x + 0 = x$ (The element 0 is called the additive identity in \mathbb{R} .)
- There is an element 1 in \mathbb{R} such that $(1)x = x$ (The element 1 is called the multiplicative identity in \mathbb{R} .)
- There is an element $-x$ in \mathbb{R} such that $x + (-x) = 0$ (The element $-x$ is the additive inverse of x in \mathbb{R} .)
- If $x \neq 0$, there is an element $\frac{1}{x}$ in \mathbb{R} such that $x(\frac{1}{x}) = 1$ (The element $\frac{1}{x}$ is the multiplicative inverse of the nonzero element x in \mathbb{R} .)
- $x(y+z) = (xy) + (xz)$ (The is the distributive property. That is, multiplication distributes over addition in \mathbb{R} .)

Any set that satisfies the properties listed in Theorem 4.2 is called a *field*. We our vectors are made from elements of a field, we call those elements of the field *scalars*.

We will algebraically represent a vector as a matrix with one column. For example, $\mathbf{v} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ is a vector with 2 entries, and we say that \mathbf{v} is a vector in 2-space. By 2-space we mean \mathbb{R}^2 , which can be geometrically modeled as the plane. Here the symbol \mathbb{R} indicates that the entries of \mathbf{v} are real numbers and the superscript 2 tells us that \mathbf{v} has two entries. Similarly, vectors in \mathbb{R}^3 have three entries, e.g., $\begin{bmatrix} 1 \\ 3 \\ -1 \end{bmatrix}$. The collection of column vectors with three entries can be geometrically modeled as three-dimensional space. If a vector \mathbf{v} has n entries we say that \mathbf{v} is a vector in \mathbb{R}^n (or n -space). Vectors are also often indicated with arrows, so we might also see a vector \mathbf{v} written as \vec{v} . It is important when writing to differentiate between a vector \mathbf{v} and a scalar v . These are quite different objects and it is up to us to make sure we are clear what a symbol represents. We will use boldface letters to represent vectors.

A vector like $\begin{bmatrix} 1 \\ 2 \end{bmatrix}$ is called a *column vector* of size 2×1 (two rows, one column). We can define an addition operation on two vectors of the same size by adding corresponding components, such as

$$\begin{bmatrix} 1 \\ -2 \end{bmatrix} + \begin{bmatrix} 3 \\ 4 \end{bmatrix} = \begin{bmatrix} 4 \\ 2 \end{bmatrix}.$$

Similarly, we can define scalar multiplication of a vector by multiplying each component of the vector by the scalar. For example,

$$3 \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 3 \\ 6 \end{bmatrix}.$$

Since we can add vectors and multiply vectors by scalars, we can then add together scalar multiples of vectors. For completeness, we define vector subtraction as adding a scalar multiple:

$$\mathbf{v} - \mathbf{u} = \mathbf{v} + (-1)\mathbf{u}.$$

This definition is equivalent to defining subtraction of \mathbf{u} from \mathbf{v} by subtracting components of \mathbf{u} from the corresponding components of \mathbf{v} .

Preview Activity 4.1.

- (1) Given vectors

$$\mathbf{v} = \begin{bmatrix} 1 \\ -2 \\ 2 \end{bmatrix}, \mathbf{u} = \begin{bmatrix} 0 \\ 1 \\ 3 \end{bmatrix}, \mathbf{w} = \begin{bmatrix} 1 \\ 1 \\ 4 \end{bmatrix},$$

determine the components of the vector $3\mathbf{v} + \mathbf{u} - 2\mathbf{w}$ using the operations defined above.

- (2) In mathematics, any time we define operations on objects, such as addition of vectors, we ask which properties the operation has. For example, one might wonder if $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$ for any two vectors \mathbf{u}, \mathbf{v} of the same size. If this property holds, we say that the *addition of vectors is a commutative operation*. However, to verify this property we cannot use examples since the property must hold for any two vectors. For simplicity, we focus on two-dimensional vectors $\mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$ and $\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$. Using these arbitrary vectors, can we say that $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$? If so, justify. If not, give a counterexample. (Note: Giving a counterexample is the best way to justify why a general statement is not true.)

- (3) One way to geometrically represent vectors with two components uses a point in the plane to correspond to a vector. Specifically, the vector $\begin{bmatrix} x \\ y \end{bmatrix}$ corresponds to the point (x, y) in the plane. As a specific example, the vector $\begin{bmatrix} 1 \\ 2 \end{bmatrix}$ corresponds to the point $(1, 2)$ in the plane. This representation will be especially handy when we consider infinite collections of vectors as we will do in this problem.

- (a) On the same set of axes, plot the points that correspond to 5-6 scalar multiples of the vector $\begin{bmatrix} 1 \\ 2 \end{bmatrix}$. Make sure to use variety of scalar multiples covering possibilities with $c > 0, c < 0, c > 1, 0 < c < 1, -1 < c < 0$. If we consider the collection of all possible scalar multiples of this vector, what do we obtain?

- (b) What would the collection of all scalar multiples of the vector $\begin{bmatrix} 0 \\ 0 \end{bmatrix}$ form in the plane?

- (c) What would the collection of all scalar multiples of the vector $\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ form in the three-dimensional space?

- (4) Let $\mathbf{u} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ and $\mathbf{v} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$ in \mathbb{R}^2 . We are interested in finding all vectors that can be formed as a sum of scalar multiples of \mathbf{u} and \mathbf{v} .



- (a) On the same set of axes, plot the points that correspond to the vectors $\mathbf{u}, \mathbf{v}, \mathbf{u} + \mathbf{v}, 1.5\mathbf{u}, 2\mathbf{v}, -\mathbf{u}, -\mathbf{v}, -\mathbf{u} + 2\mathbf{v}$. Plot other random sums of scalar multiples of \mathbf{u} and \mathbf{v} using several scalar multiples (including those less than 1 or negative) (that is, find other vectors of the form $a\mathbf{u} + b\mathbf{v}$ where a and b are any scalars.).
- (b) If we considered sums of all scalar multiples of \mathbf{u}, \mathbf{v} , which vectors will we obtain? Can we obtain any vector in \mathbb{R}^2 in this form?

Vectors and Vector Operations

As discussed in Preview Activity 4.1, a vector is simply a list of numbers. We can add vectors of like size and multiply vectors by scalars. These operations define a structure on the set of all vectors with the same number of components that will be our major object of study in linear algebra. Ultimately we will expand our idea of vectors to a more general context and study what we will call *vector spaces*.

In Preview Activity 4.1 we saw how to add vectors and multiply vectors by scalars in \mathbb{R}^2 , and this idea extends to \mathbb{R}^n for any n . Before we do so, one thing we didn't address in Preview Activity 4.1 is what it means for two vectors to be equal. It should seem reasonable that two vectors are equal if and only if they have the same corresponding components. More formally, if we let

$$\mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix} \text{ and } \mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$$

be vectors in \mathbb{R}^n , then $\mathbf{u} = \mathbf{v}$ if $u_i = v_i$ for every i between 1 and n . Note that this statement implies that a vector in \mathbb{R}^2 cannot equal a vector in \mathbb{R}^3 because they don't have the same number of components. With this in mind we can now define the sum $\mathbf{u} + \mathbf{v}$ of the vectors \mathbf{u} and \mathbf{v} to be the vector in \mathbb{R}^n defined by

$$\mathbf{u} + \mathbf{v} = \begin{bmatrix} u_1 + v_1 \\ u_2 + v_2 \\ \vdots \\ u_n + v_n \end{bmatrix}.$$

In other words, to add two vectors of the same size, we add corresponding components.

Similarly, we can define scalar multiplication of a vector. If c is a scalar, then the scalar multiple $c\mathbf{v}$ of the vector \mathbf{v} is the vector in \mathbb{R}^n defined by

$$c\mathbf{v} = \begin{bmatrix} cv_1 \\ cv_2 \\ \vdots \\ cv_n \end{bmatrix}.$$

In other words, the scalar multiple $c\mathbf{v}$ of the vector \mathbf{v} is the vector obtained by multiplying each component of the vector \mathbf{v} by the scalar c . Since we can add vectors and multiply vectors by scalars,



we can then add together scalar multiples of vectors. For completeness, we define vector subtraction as adding a scalar multiple:

$$\mathbf{v} - \mathbf{u} = \mathbf{v} + (-1)\mathbf{u}.$$

This definition is equivalent to defining subtraction of \mathbf{u} from \mathbf{v} by subtracting components of \mathbf{u} from the corresponding components of \mathbf{v} .

After defining operations on objects, we should wonder what kinds of properties these operations have. For example, with the operation of addition of real numbers we know that $1 + 2$ is equal to $2 + 1$. This is called the *commutative* property of scalar addition and says that order does not matter when we add real numbers. It is natural for us to ask if similar properties hold for the vector operations, addition and scalar multiplication, we defined. You showed in Preview Activity 4.1 that the addition operation is also commutative on vectors in \mathbb{R}^2 .

In the activity below we consider how the two operations, addition and scalar multiplication, interact with each other. In real numbers, we know that multiplication is distributive over addition. Is that true with vectors as well?

Activity 4.1. We work with vectors in \mathbb{R}^2 to make the notation easier.

Let a be an arbitrary scalar, and $\mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$ and $\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$ be two *arbitrary* vectors in \mathbb{R}^2 .

Is $a(\mathbf{u} + \mathbf{v})$ equal to $a\mathbf{u} + a\mathbf{v}$? What property does this imply about the scalar multiplication and addition operations on vectors?

Similar arguments can be used to show the following properties of vector addition and multiplication by scalars.

Theorem 4.3. Let \mathbf{v} , \mathbf{u} , and \mathbf{w} be vectors in \mathbb{R}^n and let a and b be scalars. Then

$$(1) \quad \mathbf{v} + \mathbf{u} = \mathbf{u} + \mathbf{v}$$

$$(2) \quad (\mathbf{v} + \mathbf{u}) + \mathbf{w} = \mathbf{v} + (\mathbf{u} + \mathbf{w})$$

(3) The vector $\mathbf{z} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$ has the property that $\mathbf{v} + \mathbf{z} = \mathbf{v}$. The vector \mathbf{z} is called the **zero vector**.

$$(4) \quad (-1)\mathbf{v} + \mathbf{v} = \mathbf{z}. \text{ The vector } (-1)\mathbf{v} = -\mathbf{v} \text{ is called the } \textbf{additive inverse} \text{ of the vector } \mathbf{v}.$$

$$(5) \quad (a + b)\mathbf{v} = a\mathbf{v} + b\mathbf{v}$$

$$(6) \quad a(\mathbf{v} + \mathbf{u}) = a\mathbf{v} + a\mathbf{u}$$

$$(7) \quad (ab)\mathbf{v} = a(b\mathbf{v})$$

$$(8) \quad 1\mathbf{v} = \mathbf{v}.$$

We will later see that the above properties make the set \mathbb{R}^n a *vector space*. These properties just say that, for the most part, we can manipulate vectors just as we do real numbers. Please note, though, that there is no multiplication or division of vectors.



Geometric Representation of Vectors and Vector Operations

We can geometrically represent a vector $\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$ in \mathbb{R}^2 as the point (v_1, v_2) in the plane as

we did in Preview Activity 4.1. We can similarly represent a vector $\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix}$ in \mathbb{R}^3 as the point (v_1, v_2, v_3) in the three-dimensional space. This geometric representation will be handy when we consider collections of infinitely many vectors, as we will do when we consider the span of a collection of vectors later in this section.

We can also represent the vector $\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$ in \mathbb{R}^2 as the directed line segment (or arrow) from the origin to the point (v_1, v_2) as shown in Figure 4.2 to aid in the visualization.

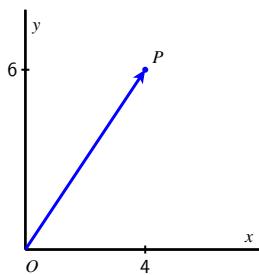


Figure 4.2: The vector $\begin{bmatrix} 4 \\ 6 \end{bmatrix}$ in \mathbb{R}^2 .

The fact that the vector in Figure 4.2 is represented by the directed line segment from the origin to the point $(4,6)$ means that this vector is the vector $\mathbf{v} = \begin{bmatrix} 4 \\ 6 \end{bmatrix}$. If O is the origin and P is the point $(4, 6)$, we will also denote this vector as \overrightarrow{OP} – so

$$\overrightarrow{OP} = \begin{bmatrix} 4 \\ 6 \end{bmatrix}.$$

In this way we can think of vectors as having direction and length. With the Pythagorean Theorem, we can see that the length of a vector $\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$ is $\sqrt{v_1^2 + v_2^2}$. This idea can be applied to vectors

in any space. If $\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \\ \vdots \\ v_n \end{bmatrix}$ is a vector in \mathbb{R}^n , then the **length** of \mathbf{v} , denoted $|\mathbf{v}|$ is the scalar

$$||\mathbf{v}|| = \sqrt{v_1^2 + v_2^2 + \cdots + v_n^2}.$$

Thinking of vectors having direction and length is especially useful in visualizing the addition of vectors. The geometric interpretation of the sum of two vectors can be seen in Figures 4.3 and 4.4.

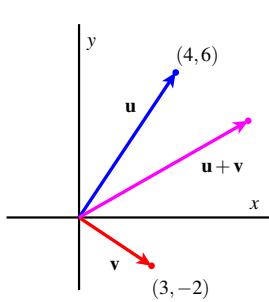


Figure 4.3: A vector sum.

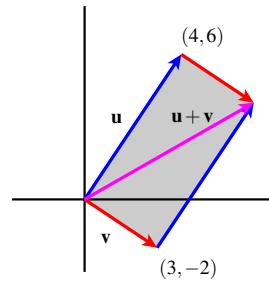


Figure 4.4: Geometric interpretation.

Let $\mathbf{u} = \begin{bmatrix} 4 \\ 6 \end{bmatrix}$ and $\mathbf{v} = \begin{bmatrix} 3 \\ -2 \end{bmatrix}$. Then $\mathbf{u} + \mathbf{v} = \begin{bmatrix} 7 \\ 4 \end{bmatrix}$ as shown in Figure 4.3. Figure 4.4 provides a context to interpret this vector sum geometrically. Using the parallelogram imposed on the three vectors, we see that if vectors \mathbf{u} and \mathbf{v} are both placed to start at the origin, then the vector sum $\mathbf{u} + \mathbf{v}$ can be visualized geometrically as the directed line segment from the origin to the fourth corner of the parallelogram.

In Preview Activity 4.1 we considered scalar multiples of a vector in \mathbb{R}^2 . The arrow representation helps in visualizing scalar multiples as well. Geometrically, a scalar multiple $c\mathbf{v}$ of a nonzero vector \mathbf{v} is a vector in the same direction as \mathbf{v} if $c > 0$ and in the opposite direction as \mathbf{v} if $c < 0$. If $c > 1$, scalar multiplication stretches the vector, while $0 < c < 1$ shrinks the vector. We also saw that the collection of all scalar multiples of a vector \mathbf{v} in \mathbb{R}^2 gives us a line through the origin and \mathbf{v} , except when $\mathbf{v} = \mathbf{0}$ in which case we only obtain $\mathbf{0}$. In other words, for a nonzero vector \mathbf{v} , the set $S = \{c\mathbf{v} : c \text{ is a scalar}\}$ is the line through the origin and \mathbf{v} in \mathbb{R}^2 .

All of these properties generalize to vectors in \mathbb{R}^3 . Specifically, the scalar multiple $c\mathbf{v}$ is a vector in the same or opposite direction as \mathbf{v} based on the sign of c , and is a stretched or shrunken version of \mathbf{v} based on whether $|c| > 1$ or $|c| < 1$. Also, the collection of all multiples of a non-zero vector \mathbf{v} in \mathbb{R}^3 form a line through the origin.

Linear Combinations of Vectors

The concept of linear combinations is one of the fundamental ideas in linear algebra. We will use linear combinations to describe almost every important concept in linear algebra – the span of a set of vectors, the range of a linear transformation, bases, the dimension of a vector space – to name just a few.

In Preview Activity 4.1, we considered the sets of all scalar multiples of a single nonzero vector in \mathbb{R}^2 and in \mathbb{R}^3 . We also considered the set of all sums of scalar multiples of two nonzero vectors. These results so far give us an idea of geometrical descriptions of sets of vectors generated by one or two vectors. Oftentimes we are interested in what vectors can be made from a given collection of vectors. For example, suppose we have two different water-benzene-acetic acid chemical solutions, one with 40% water, 50% benzene and 10% acetic acid, the other with 52% water, 42% benzene

and 6% acid. An experiment we want to conduct requires a chemical solution with 43% water, 48% benzene and 9% acid. We would like to know if we make this new chemical solution by mixing the first two chemical solutions, or do we have to run to the chemical solutions market to get the chemical solution we want.

We can set up a system of equations for each ingredient and find the answer. But we can also consider each chemical solution as a vector, where the components represent the water, benzene and acid percentages. So the two chemical solutions we have are represented by the vectors $\mathbf{v}_1 =$

$$\begin{bmatrix} 40 \\ 50 \\ 10 \end{bmatrix}$$

and $\mathbf{v}_2 = \begin{bmatrix} 52 \\ 42 \\ 6 \end{bmatrix}$. If we mix the two chemical solutions with varying amounts of each ingredient, then the question of whether we can make the desired chemical solution becomes the question of whether the equation

$$c_1 \begin{bmatrix} 40 \\ 50 \\ 10 \end{bmatrix} + c_2 \begin{bmatrix} 52 \\ 42 \\ 6 \end{bmatrix} = \begin{bmatrix} 43 \\ 48 \\ 9 \end{bmatrix}$$

has a solution. (You will determine if this equation has a solution in Exercise 5.)

We might also be interested in what other chemical solutions we can make from the two given solutions. This amounts to determining which vectors can be written in the form $c_1 \begin{bmatrix} 40 \\ 50 \\ 10 \end{bmatrix} + c_2 \begin{bmatrix} 52 \\ 42 \\ 6 \end{bmatrix}$ for scalars c_1 and c_2 . Vectors that are created from sums of scalar multiples of given vectors are called linear combinations of those vectors. More formally,

Definition 4.4. A **linear combination** of vectors $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_m$ in \mathbb{R}^n is any vector of the form

$$c_1 \mathbf{v}_1 + c_2 \mathbf{v}_2 + \cdots + c_m \mathbf{v}_m, \quad (4.1)$$

where c_1, c_2, \dots, c_m are scalars that we will refer to as the **weights**.

In the chemical solutions example, the vector $c_1 \begin{bmatrix} 40 \\ 50 \\ 10 \end{bmatrix} + c_2 \begin{bmatrix} 52 \\ 42 \\ 6 \end{bmatrix}$ for scalars c_1 and c_2 is a linear combination of the vectors $\begin{bmatrix} 40 \\ 50 \\ 10 \end{bmatrix}$ and $\begin{bmatrix} 52 \\ 42 \\ 6 \end{bmatrix}$ with weights c_1 and c_2 , and the set of linear combinations of the given chemical solution vectors tells us exactly which chemical solutions we can make from the given ones. This is one example of how linear combinations can arise in applications.

The set of all linear combinations of a fixed collection of vectors has a very nice algebraic structure and, in small dimensions, allows us to use a geometrical description to aid our understanding. In the above example, this collection gives us the type of chemical solutions we can make by combining the first two solutions in varying amounts.



Activity 4.2. Our chemical solution example illustrates that it can be of interest to determine whether certain vectors can be written as a linear combination of given vectors. We explore that

idea in more depth in this activity. Let $\mathbf{v}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ and $\mathbf{v}_2 = \begin{bmatrix} 2 \\ -1 \\ 3 \end{bmatrix}$.

- (a) Calculate the linear combination of \mathbf{v}_1 and \mathbf{v}_2 with corresponding weights (scalar multiples) 1 and 2. The resulting vector is a vector which can be written as a linear combination of \mathbf{v}_1 and \mathbf{v}_2 .

(b) Can $\mathbf{w} = \begin{bmatrix} 3 \\ 0 \\ 4 \end{bmatrix}$ be written as a linear combination of \mathbf{v}_1 and \mathbf{v}_2 ? If so, which linear combination? If not, explain why not.

(c) Can $\mathbf{w} = \begin{bmatrix} 2 \\ 0 \\ 2 \end{bmatrix}$ be written as a linear combination of \mathbf{v}_1 and \mathbf{v}_2 ? If so, which linear combination? If not, explain why not.

(d) Let $\mathbf{w} = \begin{bmatrix} 0 \\ 6 \\ -2 \end{bmatrix}$. The problem of determining if \mathbf{w} is a linear combination of \mathbf{v}_1 and \mathbf{v}_2 is equivalent to the problem of finding scalars x_1 and x_2 so that

$$\mathbf{w} = x_1\mathbf{v}_1 + x_2\mathbf{v}_2. \quad (4.2)$$

- Combine the vectors on the right hand side of equation (4.2) into one vector, and then set the components of the vectors on both sides equal to each other to convert the vector equation (4.2) to a linear system of three equations in two variables.
- Use row operations to find a solution, if it exists, to the system you found in the previous part of this activity. If you find a solution, verify in (4.2) that you have found appropriate weights to produce the vector \mathbf{w} as a linear combination of \mathbf{v}_1 and \mathbf{v}_2 .

Note that to find the weights that make \mathbf{w} a linear combination of the vectors \mathbf{v}_1 and \mathbf{v}_2 , we simply solved the linear system corresponding to the augmented matrix

$$[\mathbf{v}_1 \ \mathbf{v}_2 \ | \ \mathbf{w}],$$

where the vectors \mathbf{v}_1 , \mathbf{v}_2 , and \mathbf{w} form the columns of an augmented matrix, and the solution of the system gave us the weights of the linear combination. In general, if we want to find weights c_1, c_2, \dots, c_m so that a vector \mathbf{w} in \mathbb{R}^n is a linear combination of the vectors $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_m$ in \mathbb{R}^n , we solve the system corresponding to the augmented matrix

$$[\mathbf{v}_1 \ \mathbf{v}_2 \ \mathbf{v}_3 \ \cdots \ \mathbf{v}_m \ | \ \mathbf{w}].$$

Any solution to this system will give us the weights. If this system has no solutions, then \mathbf{w} cannot be written as a linear combination of the vectors $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_m$. This shows us the equivalence of the linear system and its vector equation representation. Specifically, we have the following result.



Theorem 4.5. *The vector equation*

$$x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + x_3\mathbf{v}_3 + \cdots + x_m\mathbf{v}_m = \mathbf{w}$$

has the same solution set as the linear system represented by the augmented matrix

$$[\mathbf{v}_1 \ \mathbf{v}_2 \ \mathbf{v}_3 \ \cdots \ \mathbf{v}_m \mid \mathbf{w}].$$

In particular, the system has a solution if and only if \mathbf{w} is a linear combination of the vectors $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \dots, \mathbf{v}_m$.

Activity 4.3.

- (a) Represent the following linear system as a vector equation. After finding the vector equation, compare your vector equation to the matrix representation you found in Preview Activity 4.1. (Note that this is the same linear system from Preview Activity 3.1.)

$$\begin{aligned} -x_3 + 3 + 2x_2 &= -x_1 \\ -3 + 2x_3 &= -x_2 \\ -2x_2 + x_1 &= 3x_3 - 7 \end{aligned}$$

- (b) Represent the following vector equation as a linear system and solve the linear system.

$$x_1 \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix} + x_2 \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} + x_3 \begin{bmatrix} -1 \\ 2 \\ -3 \end{bmatrix} = \begin{bmatrix} 4 \\ 3 \\ 11 \end{bmatrix}$$

The Span of a Set of Vectors

As we saw in the previous section, the question of whether a system of linear equations has a solution is equivalent to the question of whether the vector obtained by the non-coefficient constants in the system is a linear combination of the vectors obtained from the columns of the coefficient matrix of the system. So if we were interested in finding for which constants the system has a solution, we would look for the collection of all linear combinations of the columns. We call this collection the *span* of these vectors. In this section we investigate the concept of span both algebraically and geometrically.

Our work in Preview Activity 4.1 seems to indicate that the span of a set of vectors, i.e., the collection of all linear combinations of this set of vectors, has a nice structure. As we mentioned above, the span of a set of vectors represents the collection of all constant vectors for which a linear system has a solution, but we will also see that other important objects in linear algebra can be represented as the span of a set of vectors.

Definition 4.6. The **span** of the vectors $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_m$ in \mathbb{R}^n is the collection of all linear combinations of the vectors $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_m$.

Notation: We denote the span of a set of vectors $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_m$ as

$$\text{Span}\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_m\}.$$



So

$$\text{Span}\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_m\} = \{c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \dots + c_m\mathbf{v}_m : c_1, c_2, \dots, c_m \text{ are scalars}\}.$$

The curly braces, { }, are used in denoting sets. They represent the whole set formed by the objects included between them. So $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_m\}$ represents the collection of the vectors formed by $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_m$ for an arbitrary number m . Note that m can be 1, meaning that the collection can contain only one vector \mathbf{v}_1 .

We now investigate what the span of a set of one or two vectors is, both from an algebraic and geometric perspective, and consider what happens for more general spanning sets.

Activity 4.4.

- (a) By definition, $\text{Span}\left\{\begin{bmatrix} 1 \\ -2 \end{bmatrix}\right\}$ is the collection of all vectors which are scalar multiples of $\begin{bmatrix} 1 \\ -2 \end{bmatrix}$. Determine which vectors are in this collection. If we plot all these vectors with each vector being represented as a point in the plane, what do they form?

- (b) Let $\mathbf{v}_1 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$ and $\mathbf{v}_2 = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$ in \mathbb{R}^3 . By definition,

$$\text{Span}\left\{\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}\right\}$$

is the collection of all linear combinations of the form

$$x_1 \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} + x_2 \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix},$$

where x_1 and x_2 are any scalars.

- Find four different vectors in $\text{Span}\{\mathbf{v}_1, \mathbf{v}_2\}$ and indicate the weights (the values of x_1 and x_2) for each linear combination. (Hint: It is really easy to find 3 vectors in $\text{Span}\{\mathbf{v}_1, \mathbf{v}_2\}$ for any $\mathbf{v}_1, \mathbf{v}_2$.)

- Are there any vectors in \mathbb{R}^3 that are not in $\text{Span}\{\mathbf{v}_1, \mathbf{v}_2\}$? Explain. Verify your result.

- Set up a linear system to determine which vectors $\mathbf{w} = \begin{bmatrix} w_1 \\ w_2 \\ w_3 \end{bmatrix}$ are in $\text{Span}\{\mathbf{v}_1, \mathbf{v}_2\}$.

Specifically, which \mathbf{w} can be expressed as a linear combination of \mathbf{v}_1 and \mathbf{v}_2 ?

- Geometrically, what shape do the vectors in $\text{Span}\{\mathbf{v}_1, \mathbf{v}_2\}$ form inside \mathbb{R}^3 ?

- Is it possible for $\text{Span}\{\mathbf{z}_1, \mathbf{z}_2\}$ to be a line for two vectors $\mathbf{z}_1, \mathbf{z}_2$ in \mathbb{R}^3 ?

- What do you think are the possible geometric descriptions of a span of a set of vectors in \mathbb{R}^2 ? Explain.

- What do you think are the possible spans of a set of vectors in \mathbb{R}^3 ? Explain.



Examples

What follows are worked examples that use the concepts from this section.

Example 4.7. For each of the following systems,

- express an arbitrary solution to the system algebraically as a linear combination of vectors,
- find a set of vectors that spans the solution set,
- describe the solution set geometrically.

(a)

$$\begin{aligned} x_1 &+ x_3 = 0 \\ 2x_1 + x_2 + 3x_3 &= 0 \\ 4x_1 - x_2 + 3x_3 &= 0. \end{aligned}$$

(b)

$$\begin{aligned} x_1 + 2x_2 + 3x_3 &= 0 \\ 2x_1 + 4x_2 + 6x_3 &= 0 \\ 4x_1 + 8x_2 + 12x_3 &= 0. \end{aligned}$$

Example Solution. In each example, we use technology to find the reduced row echelon form of the augmented matrix.

(a) The reduced row echelon form of the augmented matrix

$$\left[\begin{array}{ccc|c} 1 & 0 & 1 & 0 \\ 2 & 1 & 3 & 0 \\ 4 & -1 & 3 & 0 \end{array} \right]$$

is

$$\left[\begin{array}{ccc|c} 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{array} \right].$$

- There is no pivot in the x_3 column, so x_3 is a free variable. Since the system is consistent, it has infinitely many solutions. We can write both x_1 and x_2 in terms of x_3 as $x_2 = -x_3$ and $x_1 = -x_3$. So the general solution to the system has the algebraic form

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -x_3 \\ -x_3 \\ x_3 \end{bmatrix} = x_3 \begin{bmatrix} -1 \\ -1 \\ 1 \end{bmatrix}.$$

So every solution to this system is a scalar multiple (linear combination) of the vector

$$\begin{bmatrix} -1 \\ -1 \\ 1 \end{bmatrix}.$$

- Since every solution to the system is a scalar multiple of the vector $\begin{bmatrix} -1 \\ -1 \\ 1 \end{bmatrix}$, the solution set to the system is $\text{Span} \left\{ \begin{bmatrix} -1 \\ -1 \\ 1 \end{bmatrix} \right\}$.
- As the set of scalar multiples of a single vector, the solution set to this system is a line in \mathbb{R}^3 through the origin and the point $(-1, -1, 1)$.

(b) The reduced row echelon form of the augmented matrix

$$\left[\begin{array}{ccc|c} 1 & 2 & 3 & 0 \\ 2 & 4 & 6 & 0 \\ 4 & 8 & 12 & 0 \end{array} \right]$$

is

$$\left[\begin{array}{ccc|c} 1 & 2 & 3 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array} \right].$$

- There are no pivots in the x_2 and x_3 columns, so x_2 and x_3 are free variables. Since the system is consistent, it has infinitely many solutions. We can write x_1 in terms of x_2 and x_3 as $x_1 = -2x_2 - 3x_3$. So the general solution to the system has the algebraic form

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -2x_2 - 3x_3 \\ x_2 \\ x_3 \end{bmatrix} = x_2 \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} -3 \\ 0 \\ 1 \end{bmatrix}.$$

So every solution to this system is a linear combination of the vectors $\begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix}$ and

$$\begin{bmatrix} -3 \\ 0 \\ 1 \end{bmatrix}.$$

- Since every solution to the system is a linear combination of the vectors $\begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix}$ and $\begin{bmatrix} -3 \\ 0 \\ 1 \end{bmatrix}$, the solution set to the system is

$$\text{Span} \left\{ \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -3 \\ 0 \\ 1 \end{bmatrix} \right\}.$$

- As the set of linear combinations of two vectors, the solution set to this system is a plane in \mathbb{R}^3 through the origin and the points $(-2, 1, 0)$ and $(-3, 0, 1)$.

Example 4.8. Let $W = \left\{ \begin{bmatrix} s+t \\ r+2s \\ r-3t \\ r+s+t \end{bmatrix} : r, s, t \in \mathbb{R} \right\}$.

- (a) Find three vectors \mathbf{v}_1 , \mathbf{v}_2 , and \mathbf{v}_3 such that $W = \text{Span}\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$.

(b) Can $\mathbf{w} = \begin{bmatrix} -2 \\ -4 \\ -1 \\ 0 \end{bmatrix}$ be written as a linear combination of the vectors \mathbf{v}_1 , \mathbf{v}_2 , \mathbf{v}_3 ? If so, find such a linear combination. If not, justify your response. What does your result tell us about the relationship between \mathbf{w} and W ? Explain.

(c) Can $\mathbf{u} = \begin{bmatrix} 3 \\ -4 \\ 1 \\ -1 \end{bmatrix}$ be written as a linear combination of the vectors \mathbf{v}_1 , \mathbf{v}_2 , \mathbf{v}_3 ? If so, find such a linear combination. If not, justify your response. What does your result tell us about the relationship between \mathbf{w} and W ? Explain.

- (d) What relationship, if any, exists between $\text{Span}\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ and $\text{Span } W$? Explain.

Example Solution.

- (a) Every vector in W has the form

$$\begin{bmatrix} s+t \\ r+2s \\ r-3t \\ r+s+t \end{bmatrix} = \begin{bmatrix} 0 \\ r \\ r \\ r \end{bmatrix} + \begin{bmatrix} s \\ 2s \\ 0 \\ s \end{bmatrix} + \begin{bmatrix} t \\ 0 \\ -3t \\ t \end{bmatrix} \\ = r \begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \end{bmatrix} + s \begin{bmatrix} 1 \\ 2 \\ 0 \\ 1 \end{bmatrix} + t \begin{bmatrix} 1 \\ 0 \\ -3 \\ 1 \end{bmatrix}$$

for some real numbers r , s , and t . Thus, $W = \text{Span}\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ where $\mathbf{v}_1 = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \end{bmatrix}$,

$$\mathbf{v}_2 = \begin{bmatrix} 1 \\ 2 \\ 0 \\ 1 \end{bmatrix}, \text{ and } \mathbf{v}_3 = \begin{bmatrix} 1 \\ 0 \\ -3 \\ 1 \end{bmatrix}.$$

- (b) To determine if \mathbf{w} is a linear combination of \mathbf{v}_1 , \mathbf{v}_2 , and \mathbf{v}_3 , we row reduced the augmented

matrix $[\mathbf{v}_1 \ \mathbf{v}_2 \ \mathbf{v}_3 \ | \ \mathbf{w}]$. The reduced row echelon form of the matrix $[\mathbf{v}_1 \ \mathbf{v}_2 \ \mathbf{v}_3 \ | \ \mathbf{w}]$ is

$$\left[\begin{array}{ccc|c} 1 & 0 & 0 & 2 \\ 0 & 1 & 0 & -3 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{array} \right].$$

The system with this as augmented matrix is consistent. If we let x_1 , x_2 , and x_3 be the variables corresponding to the first three columns, respectively, of this augmented matrix, then we see that $x_1 = 2$, $x_2 = -3$, and $x_3 = 1$. So \mathbf{w} can be written as a linear combination of \mathbf{v}_1 , \mathbf{v}_2 , and \mathbf{v}_3 as

$$\mathbf{w} = 2\mathbf{v}_1 - 3\mathbf{v}_2 + \mathbf{v}_3.$$

Since $W = \text{Span}\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$, it follows that $\mathbf{w} \in W$.

- (c) To determine if \mathbf{u} is a linear combination of \mathbf{v}_1 , \mathbf{v}_2 , and \mathbf{v}_3 , we row reduced the augmented matrix $[\mathbf{v}_1 \ \mathbf{v}_2 \ \mathbf{v}_3 \ | \ \mathbf{u}]$. The reduced row echelon form of the matrix $[\mathbf{v}_1 \ \mathbf{v}_2 \ \mathbf{v}_3 \ | \ \mathbf{u}]$ is

$$\left[\begin{array}{ccc|c} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{array} \right].$$

The last row shows that the system with this as augmented matrix is inconsistent. So \mathbf{u} cannot be written as a linear combination of \mathbf{v}_1 , \mathbf{v}_2 , and \mathbf{v}_3 . Since $W = \text{Span}\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$, it follows that $\mathbf{u} \notin W$.

- (d) We know that $\text{Span}\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\} = W$. Now $\text{Span } W$ contains the linear combinations of vectors in W , which are all linear combinations of the vectors \mathbf{v}_1 , \mathbf{v}_2 , and \mathbf{v}_3 . Thus, $\text{Span } W$ is just the set of linear combinations of \mathbf{v}_1 , \mathbf{v}_2 , and \mathbf{v}_3 . We conclude that $\text{Span } W = \text{Span}\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\} = W$.

Summary

- A vector is a list of numbers in a specified order.
- We add two vectors of the same size by adding corresponding components. In other words, if \mathbf{u} and \mathbf{v} are vectors of the same size and u_i and v_i are the i components of \mathbf{u} and \mathbf{v} , respectively, then $\mathbf{u} + \mathbf{v}$ is the vector whose i th component is $u_i + v_i$ for each i . Geometrically, we represent the sum of two vectors using the Parallelogram Rule: The vector $\mathbf{u} + \mathbf{v}$ is the directed line segment from the origin to the 4th point of the parallelogram formed by the origin and the vectors \mathbf{u}, \mathbf{v} .
- A scalar multiple of a vector is found by multiplying each component of the vector by that scalar. In other words, if v_i is the i component of the vector \mathbf{v} and c is any scalar, then $c\mathbf{v}$ is the vector whose i component is cv_i for each i . Geometrically, a scalar multiple of a nonzero vector \mathbf{v} is a vector in the same direction as \mathbf{v} if $c > 0$ and in the opposite direction if $c < 0$. If $|c| > 1$, the vector is stretched, and if $|c| < 1$, the vector is shrunk.



- An important concept is that of a linear combination of vectors. In words, a linear combination of a collection of vectors is a sum of scalar multiples of the vectors. More formally, we defined a linear combination of vectors $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_m$ in \mathbb{R}^n is any vector of the form $c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \dots + c_m\mathbf{v}_m$, where c_1, c_2, \dots, c_m are scalars.
- To find weights c_1, c_2, \dots, c_m so that a vector \mathbf{w} in \mathbb{R}^n is a linear combination of the vectors $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_m$ in \mathbb{R}^n , we simply solve the system corresponding to the augmented matrix

$$[\mathbf{v}_1 \ \mathbf{v}_2 \ \mathbf{v}_3 \ \cdots \ \mathbf{v}_m \mid \mathbf{w}].$$

- The collection of all linear combinations of a set of vectors is called the span of the set of vectors. More formally, the span of the vectors $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_m$ in \mathbb{R}^n is the set

$$\{c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \dots + c_m\mathbf{v}_m : c_1, c_2, \dots, c_m \text{ are scalars}\},$$

which we denote as $\text{Span}\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_m\}$. Geometrically, the span of a single nonzero vector \mathbf{v} in any dimension is the line through the origin and the vector \mathbf{v} . The span of two vectors $\mathbf{v}_1, \mathbf{v}_2$ in any dimension neither of which is a multiple of the other is a plane through the origin containing both vectors.

Exercises

- (1) Given vectors $\mathbf{u} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ and $\mathbf{v} = \begin{bmatrix} -1 \\ 2 \end{bmatrix}$ in \mathbb{R}^2 , determine if $\mathbf{w} = \begin{bmatrix} -4 \\ -1 \end{bmatrix}$ can be written as a linear combination of \mathbf{u} and \mathbf{v} . If so, determine the weights of \mathbf{u} and \mathbf{v} which produce \mathbf{w} .

- (2) Given vectors $\mathbf{v}_1 = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}$, $\mathbf{v}_2 = \begin{bmatrix} -2 \\ 1 \\ 2 \end{bmatrix}$ and $\mathbf{v}_3 = \begin{bmatrix} -1 \\ 3 \\ 3 \end{bmatrix}$ in \mathbb{R}^3 , determine if $\mathbf{w} = \begin{bmatrix} 5 \\ 5 \\ 1 \end{bmatrix}$ can be written as a linear combination of $\mathbf{v}_1, \mathbf{v}_2$ and \mathbf{v}_3 . If so, determine the weights of $\mathbf{v}_1, \mathbf{v}_2$ and \mathbf{v}_3 which produce \mathbf{w} . Reflect on the result. Is there anything special about the given vectors $\mathbf{v}_1, \mathbf{v}_2$ and \mathbf{v}_3 ?

- (3) Let $\mathbf{u} = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}$ and $\mathbf{v} = \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix}$ in \mathbb{R}^3 . Determine which vectors $\mathbf{w} = \begin{bmatrix} w_1 \\ w_2 \\ w_3 \end{bmatrix}$ in \mathbb{R}^3 can be written as a linear combination of \mathbf{u} and \mathbf{v} . Does the set of \mathbf{w} 's include the $\mathbf{0}$ vector? If so, determine which weights in the linear combination produce the $\mathbf{0}$ vector. If not, explain why not.

- (4) Consider vectors $\mathbf{u} = \begin{bmatrix} 0 \\ 2 \\ 0 \end{bmatrix}$ and $\mathbf{v} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ in \mathbb{R}^3 .

(a) Find four specific linear combinations of the vectors \mathbf{u} and \mathbf{v} .

(b) Explain why the zero vector must be a linear combination of \mathbf{u} and \mathbf{v} .



- (c) What kind of geometric shape does the set of all linear combinations of \mathbf{u} and \mathbf{v} have in \mathbb{R}^3 ?
- (d) Can we obtain any vector in \mathbb{R}^3 as a linear combination of \mathbf{u} and \mathbf{v} ? Explain.
- (5) Suppose we have two different water-benzene-acetic acid solutions, one with 40% water, 50% benzene and 10% acetic acid, the other with 52% water, 42% benzene and 6% acid.
- An experiment we want to conduct requires a solution with 43% water, 48% benzene and 9% acid. Representing each acid solution as a vector, determine if we can we make this new acid solution by mixing the first two solutions, or do we have to run to the chemical solutions market to get the solution we want?
 - Using the water-benzene-acetic acid solutions in the previous problem, can we obtain an acid solution which contains 50% water, 43% benzene and 7% acid?
 - Determine the relationship between the percentages of water, benzene, and acid in solutions which can be obtained by mixing the two given water-benzene-acetic acid solutions above.
- (6) Is the vector $\mathbf{b} = \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix}$ in $\text{Span} \left\{ \begin{bmatrix} 2 \\ -1 \\ 0 \end{bmatrix}, \begin{bmatrix} -3 \\ 0 \\ -5 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} \right\}$? Justify your answer.
- (7) Describe geometrically each of the following sets.
- $\text{Span} \left\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ -1 \end{bmatrix} \right\}$ in \mathbb{R}^2
 - $\text{Span} \left\{ \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ -1 \\ -1 \end{bmatrix}, \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix} \right\}$ in \mathbb{R}^3
- (8) Consider the linear system
- $$\begin{aligned} 2x_1 + 3x_2 + 3x_3 &= 0 \\ 4x_1 + 6x_3 + 6x_4 &= 0 \\ 2x_1 + 4x_2 + 3x_3 - x_4 &= 0. \end{aligned}$$
- (a) Find the general solution to this system.
- (b) Find two specific vectors \mathbf{v}_1 and \mathbf{v}_2 so that the solution set to this system is $\text{Span}\{\mathbf{v}_1, \mathbf{v}_2\}$.
- (9) Answer the following question as yes or no. Verify your answer. If \mathbf{u} and \mathbf{v} are vectors in \mathbb{R}^n , then \mathbf{v} is in $\text{Span}\{\mathbf{u}, \mathbf{u} - \mathbf{v}\}$.
- (10) Let \mathbf{v} , \mathbf{u} , and \mathbf{w} be vectors in \mathbb{R}^n and let a and b be scalars. Verify Theorem 4.3. That is, show that
- $\mathbf{v} + \mathbf{u} = \mathbf{u} + \mathbf{v}$
 - $(\mathbf{v} + \mathbf{u}) + \mathbf{w} = \mathbf{v} + (\mathbf{u} + \mathbf{w})$



- (c) The vector $\mathbf{z} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$ has the property that $\mathbf{v} + \mathbf{z} = \mathbf{v}$.
- (d) $(-1)\mathbf{v} + \mathbf{v} = \mathbf{z}$.
- (e) $(a+b)\mathbf{v} = a\mathbf{v} + b\mathbf{v}$
- (f) $a(\mathbf{v} + \mathbf{u}) = a\mathbf{v} + a\mathbf{u}$
- (g) $(ab)\mathbf{v} = a(b\mathbf{v})$
- (h) $1\mathbf{v} = \mathbf{v}$.
- (11) Label each of the following statements as True or False. Provide justification for your response.
- True/False** A vector in \mathbb{R}^2 , i.e. a two-dimensional vector, is also a vector in \mathbb{R}^3 .
 - True/False** Any vector in \mathbb{R}^2 can be visualized as a vector in \mathbb{R}^3 by adding a 0 as the last coordinate.
 - True/False** The zero vector is a scalar multiple of any other vector (of the same size).
 - True/False** The zero vector cannot be a linear combination of two non-zero vectors.
 - True/False** Given two vectors \mathbf{u} and \mathbf{v} , the vector $\frac{1}{2}\mathbf{u}$ is a linear combination of \mathbf{u} and \mathbf{v} .
 - True/False** Given any two non-zero vectors \mathbf{u} and \mathbf{v} in \mathbb{R}^2 , we can obtain any vector in \mathbb{R}^2 as a linear combination of \mathbf{u} and \mathbf{v} .
 - True/False** Given any two distinct vectors \mathbf{u} and \mathbf{v} in \mathbb{R}^2 , we can obtain any vector in \mathbb{R}^2 as a linear combination of \mathbf{u} and \mathbf{v} .
 - True/False** If \mathbf{u} can be expressed as a linear combination of \mathbf{v}_1 and \mathbf{v}_2 , then $2\mathbf{u}$ can also be expressed as a linear combination of \mathbf{v}_1 and \mathbf{v}_2 .
 - True/False** The span of any two vectors neither of which is a multiple of the other can be visualized as a plane through the origin.
 - True/False** Given any vector, the collection of all linear combinations of this vector can be visualized as a line through the origin.
 - True/False** The span of any collection of vectors includes the $\mathbf{0}$ vector.
 - True/False** If the span of \mathbf{v}_1 and \mathbf{v}_2 is all of \mathbb{R}^2 , then so is the span of \mathbf{v}_1 and $\mathbf{v}_1 + \mathbf{v}_2$.
 - True/False** If the span of \mathbf{v}_1 , \mathbf{v}_2 and \mathbf{v}_3 is all of \mathbb{R}^3 , then so is the span of $\mathbf{v}_1 + \mathbf{v}_2$ and $\mathbf{v}_2 + \mathbf{v}_3$.

Project: Analyzing Knight Moves

To understand where a knight can move in a chess game, we need to know the initial setup. A chess board is an 8×8 grid. To be able to refer to the individual positions on the board, we will place the



board so that its lower left corner is at the origin, make each square in the grid have side length 1, and label each square with the point at the lower left corner. This is illustrated at left in Figure 4.5.

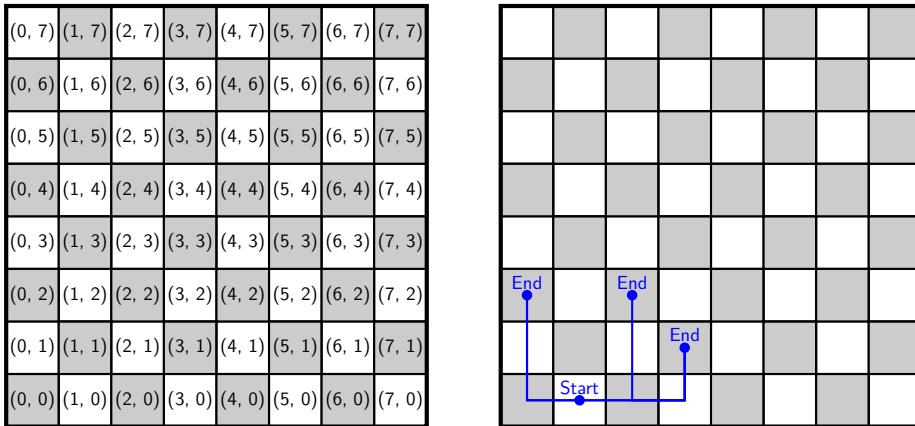


Figure 4.5: Initial knight placement and moves.

Each player has two knights to start the game, for one player the knights would begin in positions $(1, 0)$ and $(6, 0)$. Because of the symmetry of the knight's moves, we will only analyze the moves of the knight that begins at position $(1, 0)$. This knight has only three allowable moves from its starting point (assuming that the board is empty), as shown at right in Figure 4.5. The questions we will ask are: given any position on the board, can the knight move from its start position to that position using only knight moves and, what sequence of moves will make that happen. To answer these questions we will use linear combinations of knight moves described as vectors.

Each knight move can be described by a vector. A move one position to the right and two up can be represented as $\mathbf{n}_1 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$. Three other moves are $\mathbf{n}_2 = \begin{bmatrix} -1 \\ 2 \end{bmatrix}$, $\mathbf{n}_3 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$, and $\mathbf{n}_4 = \begin{bmatrix} -2 \\ 1 \end{bmatrix}$. The other four knight moves are the additive inverses of these four. Any sequence of moves by the knight is given by the linear combination

$$x_1\mathbf{n}_1 + x_2\mathbf{n}_2 + x_3\mathbf{n}_3 + x_4\mathbf{n}_4.$$

A word of caution: the knight can only make complete moves, so we are restricted to integer (either positive or negative) values for x_1 , x_2 , x_3 , and x_4 . You can use the GeoGebra app at <https://www.geogebra.org/m/dfwtskrj> to see the effects the weights have on the knight moves. We should note here that since addition of vectors is commutative, the order in which we apply our moves does not matter. However, we may need to be careful with the order so that our knight does not leave the chess board.

Project Activity 4.1.

- (a) Explain why the vector equation

$$\begin{bmatrix} 1 \\ 0 \end{bmatrix} + x_1\mathbf{n}_1 + x_2\mathbf{n}_2 + x_3\mathbf{n}_3 + x_4\mathbf{n}_4 = \begin{bmatrix} 5 \\ 2 \end{bmatrix}$$

will tell us if it is possible for the knight to move from its initial position at $(1, 0)$ to the position $(5, 2)$.

- (b) Find all solutions, if any, to the system from part (a). If it is possible to find a sequence of moves that take the knight from its initial position to position $(5, 2)$, find weights x_1 , x_2 , x_3 , and x_4 to accomplish this move. (Be careful – we must have solutions in which x_1 , x_2 , x_3 , and x_4 are integers.) Is there more than one sequence of possible moves? You can check your solution with the GeoGebra app at <https://www.geogebra.org/m/dfwtskrj>.

Project Activity 4.1 shows that it is possible for our knight to move to position $(5, 2)$ on the board. We would like to know if it is possible to move to any position on the board. That is, we would like to know if the integer span of the four moves \mathbf{n}_1 , \mathbf{n}_2 , \mathbf{n}_3 , and \mathbf{n}_4 will allow our knight to cover the entire board. This takes a bit more work.

Project Activity 4.2. Given any position (a, b) , we want to know if our knight can move from its start position $(1, 0)$ to position (a, b) .

- (a) Write a vector equation whose solution will tell us if it is possible for our knight to move from its start position $(1, 0)$ to position (a, b) .
 (b) Show that the solution to part (a) can be written in the form

$$x_1 = \frac{1}{4}(-5x_3 + 3x_4 + b + 2(a - 1)) \quad (4.3)$$

$$x_2 = \frac{1}{4}(3x_3 - 5x_4 + b - 2(a - 1)) \quad (4.4)$$

x_3 is free

x_4 is free.

To answer our question if our knight can reach any position, we now need to determine if we can always find integer values of x_3 and x_4 to make equations (4.3) and (4.4) have integer solutions. In other words, we need to find values of x_3 and x_4 so that $-5x_3 + 3x_4 + b + 2(a - 1)$ and $3x_3 - 5x_4 + b - 2(a - 1)$ are multiples of 4. How we do this could depend on the parity (even or odd) of a and b . For example, if a is odd and b is even, say $a = 2r + 1$ and $b = 2s$ for some integers r and s , then

$$x_1 = \frac{1}{4}(-5x_3 + 3x_4 + 2s + 4r)$$

$$x_2 = \frac{1}{4}(3x_3 - 5x_4 + 2s - 4r).$$

With a little trial and error we can see that if we let $x_3 = x_4 = s$, then $x_1 = r$ and $x_2 = -r$ is a solution with integer weights. For example, when $a = 5$ and $b = 2$ we have $r = 2$ and $s = 1$. This makes $x_1 = 2$, $x_2 = -2$, $x_3 = 1 = x_4$. Compare this to the solution(s) you found in Project Activity 4.1. This analysis shows us how to move our knight to any position (a, b) where a is odd and b is even.

Project Activity 4.3. Complete the analysis as above to determine if there are integer solutions to our knight's move system in the following cases.



- (a) a odd and b odd
- (b) a even and b even
- (c) a even and b odd.

Project Activity 4.3 shows that for any position on the chess board, using linear combinations of move vectors, we can find a sequence of moves that takes our knight to that position. (We actually haven't shown that these moves can be made so that our knight always stays on the board – we leave that question to you.)

Section 5

The Matrix-Vector Form of a Linear System

Focus Questions

By the end of this section, you should be able to give precise and thorough answers to the questions listed below. You may want to keep these questions in mind to focus your thoughts as you complete the section.

- How and when is the matrix-vector product Ax defined?
- How can a system of linear equations be written in matrix-vector form?
- How can we tell if the system $Ax = b$ is consistent for a given vector b ?
- How can we tell if the system $Ax = b$ is consistent for every vector b ?
- What is a homogeneous system? What can we say about the solution set to a homogeneous system?
- What must be true about pivots in the coefficient matrix A in order for the homogeneous system $Ax = \mathbf{0}$ to have a unique solution?
- How are the solutions to the nonhomogeneous system $Ax = b$ related to the solutions of the corresponding homogeneous system $Ax = \mathbf{0}$?

Application: Modeling an Economy

An economy is a very complex system. An economy is not a well-defined object, there are many factors that influence an economy, and it is often unclear how the factors influence each other. Mathematical modeling plays an important role in attempting to understand an economy.

In 1941 Wassily Leontief developed the first empirical model of a national economy. Around 1949 Leontief used data from the U.S. Bureau of Labor Statistics to divide the U.S. economy into

500 sectors. He then set up linear equations for each sector. This system was too large for the computers at the time to solve, so he then aggregated the information into 42 sectors. The Harvard Mark II computer was used to solve this system, one of the first significant uses of computers for mathematical modeling. Leontief won the 1973 Nobel Prize in economics for his work.

With such large models (Leontief's models are called *input-output* models) it is important to find a shorthand way to represent the resulting systems. In this section we will see how to represent any size system of linear equations in a very convenient way. Later, we will analyze a small economy using input-output models.

Introduction

There is another useful way to represent a system of linear equations using a matrix-vector product that we investigate in this section. To understand how this product comes about, recall that we can represent the linear system

$$\begin{aligned}x_1 + 4x_2 + 2x_3 + 4x_4 &= 1 \\2x_1 - x_2 - 5x_3 - x_4 &= 2 \\3x_1 + 7x_2 + x_3 + 7x_4 &= 3\end{aligned}$$

as a vector equation as

$$x_1 \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} + x_2 \begin{bmatrix} 4 \\ -1 \\ 7 \end{bmatrix} + x_3 \begin{bmatrix} 2 \\ -5 \\ 1 \end{bmatrix} + x_4 \begin{bmatrix} 4 \\ -1 \\ 7 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}. \quad (5.1)$$

We can view the left hand side of Equation (5.1) as a **matrix-vector product**. Specifically, if

$$A = \begin{bmatrix} 1 & 4 & 2 & 4 \\ 2 & -1 & -5 & -1 \\ 3 & 7 & 1 & 7 \end{bmatrix} \text{ and } \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}, \text{ then we define the } \textit{matrix-vector product} A\mathbf{x} \text{ as}$$

the left hand side Equation (5.1). So the matrix vector product $A\mathbf{x}$ is the linear combination of the columns of A with weights from the vector \mathbf{x} in order.

With this definition, the vector equation in (5.1) can be expressed as a matrix-vector equation as

$$\begin{bmatrix} 1 & 4 & 2 & 4 \\ 2 & -1 & -5 & -1 \\ 3 & 7 & 1 & 7 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}.$$

We call this representation the **matrix-vector form** of the system. Note that the matrix A in this expression is the same as the coefficient matrix that appears in the augmented matrix representation of the system.

We can use the above definition of the matrix-vector product as a linear combination with any matrix and any vector, as long as it is meaningful to use the entries in the vector as weights for the

columns of the matrix. For example, for $A = \begin{bmatrix} 1 & 2 \\ 3 & 1 \\ 1 & 1 \end{bmatrix}$ and $\mathbf{v} = \begin{bmatrix} 3 \\ 4 \end{bmatrix}$, then we can define $A\mathbf{v}$ to be the linear combination of the columns of A with weights 3 and 4:

$$A\mathbf{v} = 3 \begin{bmatrix} 1 \\ 3 \\ 1 \end{bmatrix} + 4 \begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 11 \\ 13 \\ 7 \end{bmatrix}.$$

However, note that if \mathbf{v} had three entries, this definition would not make sense since we do not have three columns in A . In those cases, we say $A\mathbf{v}$ is not defined. We will later see that this definition can be generalized to matrix-matrix products, by treating the vector as a special case of a matrix with one column.

Preview Activity 5.1.

- (1) Write the vector equation

$$x_1 \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix} + x_2 \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} + x_3 \begin{bmatrix} -1 \\ 2 \\ -3 \end{bmatrix} = \begin{bmatrix} 4 \\ 3 \\ 11 \end{bmatrix}$$

in matrix-vector form. Note that this is the vector equation whose augmented matrix representation was given in Problem 2 in Preview Activity 2.1. Compare your matrix A and the right hand side vector to the augmented matrix. Do not solve the system.

- (2) Given the matrix-vector equation

$$\begin{bmatrix} 1 & 2 & -1 \\ 0 & 1 & 2 \\ 1 & -2 & -3 \end{bmatrix} \mathbf{x} = \begin{bmatrix} -3 \\ 3 \\ -7 \end{bmatrix}$$

represent the system corresponding to this equation. Note that this should correspond to the system (or an equivalent system where an equation might be multiplied by (-1)) in Problem 1 of Preview Activity 2.1.

- (3) Find the indicated matrix-vector products, if possible. Express as one vector.

(a) $\begin{bmatrix} 2 & -2 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \end{bmatrix}$

(b) $\begin{bmatrix} 1 & 0 & 2 \\ 2 & -2 & 3 \end{bmatrix} \begin{bmatrix} 2 \\ 1 \end{bmatrix}$

(c) $\begin{bmatrix} -6 & -2 & 1 \\ 2 & -2 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ -3 \\ 1 \end{bmatrix}$

- (4) As you might have noticed, systems with all the constants being 0 are special in that they always have a solution. (Why?) So we might consider grouping systems into two types: Those of the form $A\mathbf{x} = \mathbf{b}$, where not all of the entries of the vector \mathbf{b} are 0, and those of the form $A\mathbf{x} = \mathbf{0}$, where $\mathbf{0}$ is the vector of all zeros. Systems like $A\mathbf{x} = \mathbf{b}$, where \mathbf{b} contains

at least one non-zero entry, are called *nonhomogeneous* systems, and systems of the form $A\mathbf{x} = \mathbf{0}$ are called *homogeneous* systems. For every nonhomogeneous system $A\mathbf{x} = \mathbf{b}$ there is a corresponding homogeneous system $A\mathbf{x} = \mathbf{0}$, and there is a useful connection between the solutions to the nonhomogeneous system and the corresponding homogeneous system. For example, consider the nonhomogeneous system

$$A\mathbf{x} = \mathbf{b}$$

with

$$A = \begin{bmatrix} 1 & 1 & 2 \\ 1 & 2 & 1 \end{bmatrix}, \quad \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}, \quad \text{and} \quad \mathbf{b} = \begin{bmatrix} 0 \\ -2 \end{bmatrix}. \quad (5.2)$$

The augmented matrix representation of this system is $[A | \mathbf{b}]$. If we reduce this augmented matrix, we find

$$\left[\begin{array}{ccc|c} 1 & 0 & 3 & 2 \\ 0 & 1 & -1 & -2 \end{array} \right].$$

From this RREF, we immediately see that the general solution is that x_3 is free, $x_2 = x_3 - 2$, and $x_1 = 2 - 3x_3$. In vector form, we can represent this general solution as

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 2 - 3x_3 \\ x_3 - 2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 2 \\ -2 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} -3 \\ 1 \\ 1 \end{bmatrix}. \quad (5.3)$$

The rightmost expression above is called the *parametric vector form* of the solution.

If we had a system where the general solution involved more than one free variable, then we would write the parametric vector form to include one vector multiplying each free variable. For example, if the general solution of a system were that x_2 and x_3 are free and $x_1 = 2 + x_2 + 3x_3$, then the parametric vector form would be

$$\mathbf{x} = \begin{bmatrix} 2 + x_2 + 3x_3 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 2 \\ 0 \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} 3 \\ 0 \\ 1 \end{bmatrix}.$$

Note that the parametric vector form expresses the solutions as a linear combination of a number of vectors, depending on the number of free variables, with an added constant vector. This expression helps us in interpreting the solution set geometrically, as we will see in this section.

- (a) Find the general solution to the homogeneous system

$$A\mathbf{x} = \mathbf{0}$$

with A and \mathbf{x} as in (5.2) and compare it to the solution to the nonhomogeneous system in (5.3). What do you notice?

- (b) Find the general solution to the nonhomogeneous system

$$A\mathbf{x} = \mathbf{b}$$



with

$$A = \begin{bmatrix} 1 & 2 & -1 \\ 2 & 4 & -2 \end{bmatrix}, \quad \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}, \quad \text{and} \quad \mathbf{b} = \begin{bmatrix} -1 \\ 1 \end{bmatrix}.$$

and express it in parametric vector form. Then find the general solution to the corresponding homogeneous system and express it in parametric vector form. How are the two solution sets related?

- (c) Make a conjecture about the relationship between the solutions to a consistent non-homogeneous system $A\mathbf{x} = \mathbf{b}$ and the corresponding homogeneous system $A\mathbf{x} = \mathbf{0}$. Be as specific as possible.

The Matrix-Vector Product

The matrix-vector product we defined in Preview Activity 5.1 for a specific example generalizes in a very straightforward manner, and provides a convenient way to represent a system of linear equations of any size using matrices and vectors. In addition to providing us with an algebraic approach to solving systems via matrices and vectors – leading to a powerful geometric relationship between solution sets of homogeneous and non-homogeneous systems – this representation allows us to think of a linear system from a dynamic perspective, as we will see later in the section on matrix transformations.

The matrix-vector product $A\mathbf{x}$ is a linear combination of the columns of A with weights from \mathbf{x} . To define this product in general, we will need a little notation. Recall that a matrix is made of rows and columns – the entries reading from left to right form the *rows* of the matrix and the entries reading from top to bottom form the *columns*. For example, the matrix

$$A = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \end{bmatrix}.$$

has three rows and four columns. The number of rows and columns of a matrix is called the *size* of the matrix, so A is a 3 by 4 matrix (also written as 3×4). We often need to have a way to reference the individual entries of a matrix A , and to do so we typically give a label, say a_{ij} to the entry in the i th row and j th column of A . So in our example we have $a_{23} = 7$. We also write $A = [a_{ij}]$ to indicate a matrix whose i, j th entry is a_{ij} . At times it is convenient to write a matrix in terms of its rows or columns. If $A = [a_{ij}]$ is an $m \times n$ matrix, then we will write

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n-1} & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n-1} & a_{2n} \\ \vdots & & \ddots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn-1} & a_{mn} \end{bmatrix}$$



or, if we let $\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_m$ denote the rows of the matrix A , then we can write A as¹

$$A = \begin{bmatrix} \mathbf{r}_1 \\ \mathbf{r}_2 \\ \vdots \\ \mathbf{r}_m \end{bmatrix}.$$

We can also write A in terms of its columns, $\mathbf{c}_1, \mathbf{c}_2, \dots, \mathbf{c}_n$, as

$$A = [\mathbf{c}_1 \ \mathbf{c}_2 \ \cdots \ \mathbf{c}_n].$$

In general, the product of a matrix with a vector is defined as follows.

Definition 5.1. Let A be an $m \times n$ matrix with columns $\mathbf{c}_1, \mathbf{c}_2, \dots, \mathbf{c}_n$, and let $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$ be a vector in \mathbb{R}^n . The **matrix-vector product** $A\mathbf{x}$ is

$$A\mathbf{x} = x_1\mathbf{c}_1 + x_2\mathbf{c}_2 + \cdots + x_n\mathbf{c}_n.$$

Important Note: The matrix-vector product $A\mathbf{x}$ is defined only when the number of entries of the vector \mathbf{x} is equal to the number of columns of the matrix A . That is, if A is an $m \times n$ matrix, then $A\mathbf{x}$ is defined only if \mathbf{x} is a column vector with n entries.

The Matrix-Vector Form of a Linear System

As we saw in Preview Activity 5.1, the matrix-vector product provides us with a short hand way of representing a system of linear equations. In general, every linear system can be written in matrix-vector form as follows.

The linear system

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\ &\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_m \end{aligned}$$

of m equations in n unknowns can be written in matrix-vector form as $A\mathbf{x} = \mathbf{b}$, where

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}, \quad \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, \quad \text{and} \quad \mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}.$$

¹Technically, the rows of A are made from the entries of the row vectors, but we use this notation as a shorthand.



This general system can also be written in the vector form

$$x_1 \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix} + x_2 \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{bmatrix} + \cdots + x_n \begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}.$$

With this last representation, we now have four different ways to represent a system of linear equations (as a system of linear equations, as an augmented matrix, in vector equation form, and in matrix-vector equation form), and it is important to be able to translate between them. As an example, the system

$$\begin{aligned} x_1 + 4x_2 + 2x_3 + 4x_4 &= 2 \\ 2x_1 - x_2 - 5x_3 - x_4 &= 2 \\ 3x_1 + 7x_2 + x_3 + 7x_4 &= 3 \end{aligned}$$

from the introduction to this section has corresponding augmented matrix

$$\left[\begin{array}{cccc|c} 1 & 4 & 2 & 4 & 1 \\ 2 & -1 & -5 & -1 & 2 \\ 3 & 7 & 1 & 7 & 3 \end{array} \right],$$

is expressed in vector form as

$$x_1 \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} + x_2 \begin{bmatrix} 4 \\ -1 \\ 7 \end{bmatrix} + x_3 \begin{bmatrix} 2 \\ -5 \\ 1 \end{bmatrix} + x_4 \begin{bmatrix} 4 \\ -1 \\ 7 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix},$$

and has matrix-vector form

$$\left[\begin{array}{cccc} 1 & 4 & 2 & 4 \\ 2 & -1 & -5 & -1 \\ 3 & 7 & 1 & 7 \end{array} \right] \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}.$$

Activity 5.1. In this activity, we will use the equivalence of the different representations of a system to make useful observations about when a system represented as $\mathbf{A}\mathbf{x} = \mathbf{b}$ has a solution.

- (a) Consider the system

$$\left[\begin{array}{ccc} 1 & 2 & -1 \\ 2 & 1 & 3 \end{array} \right] \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 2 \\ 6 \end{bmatrix}.$$

Write the matrix-vector product on the left side of this equation as a linear combination of the columns of the coefficient matrix. Find weights that make the vector $\begin{bmatrix} 2 \\ 6 \end{bmatrix}$ a linear combination of the columns of the coefficient matrix.

- (b) From this point on we consider the general case where A is an $m \times n$ matrix. Use the vector equation representation to explain why the system $Ax = b$ has a solution if and only if b is a linear combination of the columns of A . (Note that ‘if and only if’ is an expression to mean that if one side of the expression is true, then the other side must also be true.) (Hint: Compare to what you did in part (a).)
- (c) Use part (b) and the definition of span to explain why the system $Ax = b$ has a solution if and only if the vector b is in the span of the columns of A .
- (d) Use part (c) to explain why the system $Ax = b$ always has a solution for any vector b in \mathbb{R}^m if and only if the span of the columns of A is all of \mathbb{R}^m .
- (e) Use the augmented matrix representation and the criterion for a consistent system to explain why the system $Ax = b$ is consistent for all vectors b if and only if A has a pivot position in every row.

We summarize our observations from the above activity in the following theorem.

Theorem 5.2. *Let A be an $m \times n$ matrix. The following statements are equivalent:*

- (1) *The matrix equation $Ax = b$ has a solution for every vector b in \mathbb{R}^m .*
- (2) *Every vector b in \mathbb{R}^m can be written as a linear combination of the columns of A .*
- (3) *The span of the columns of A is \mathbb{R}^m .*
- (4) *The matrix A has a pivot position in each row.*

In the future, if we need to determine whether a system has a solution for every b , we can refer to this theorem without having to argue our reasoning from scratch.

Properties of the Matrix Vector Product

As we have done before, we have a new operation (the matrix-vector product), so we should wonder what properties it has.

Activity 5.2. In this activity, we consider whether the matrix-vector product distributes vector addition. In other words: Is $A(\mathbf{u} + \mathbf{v})$ equal to $A\mathbf{u} + A\mathbf{v}$?

We work with arbitrary vectors \mathbf{u}, \mathbf{v} in \mathbb{R}^3 and an arbitrary matrix A with 3 columns (so that $A\mathbf{u}$ and $A\mathbf{v}$ are defined) to simplify notation. Let $A = [\mathbf{c}_1 \mathbf{c}_2 \mathbf{c}_3]$ (note that each \mathbf{c}_i represents a column of A), $\mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix}$, and $\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix}$. Use the definition of the matrix-vector product along with the properties of vector operations to show that

$$A(\mathbf{u} + \mathbf{v}) = A\mathbf{u} + A\mathbf{v}.$$

Similar arguments using the definition of matrix-vector product along with the properties of vector operations can be used to show the following theorem:



Theorem 5.3. Let A be an $m \times n$ matrix, \mathbf{u} and \mathbf{v} $n \times 1$ vectors, and c a scalar. Then

$$(1) \quad A(\mathbf{u} + \mathbf{v}) = A\mathbf{u} + A\mathbf{v}$$

$$(2) \quad c(A\mathbf{v}) = A(c\mathbf{v})$$

Homogeneous and Nonhomogeneous Systems

As we saw before, the systems with all the right hand side constants being 0 are special in that they always have a solution. (Why?) So we might consider grouping systems into two types: Those of the form $A\mathbf{x} = \mathbf{b}$, where not all of the entries of the vector \mathbf{b} are 0, and those of the form $A\mathbf{x} = \mathbf{0}$, where $\mathbf{0}$ is the vector of all zeros. Systems like $A\mathbf{x} = \mathbf{b}$, where \mathbf{b} contains at least one non-zero entry, are called **nonhomogeneous** systems, and systems of the form $A\mathbf{x} = \mathbf{0}$ are called **homogeneous** systems. For every nonhomogeneous system $A\mathbf{x} = \mathbf{b}$ there is a corresponding homogeneous system $A\mathbf{x} = \mathbf{0}$. We now investigate the connection between the solutions to the nonhomogeneous system and the corresponding homogeneous system.

Activity 5.3. In this activity we will consider the relationship between the solution sets of nonhomogeneous systems and those of the corresponding homogeneous systems.

- (a) Find the solution sets of the system

$$A\mathbf{x} = \mathbf{b}$$

where

$$A = \begin{bmatrix} 1 & 1 & 2 \\ 1 & 2 & 1 \end{bmatrix}, \quad \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}, \quad \text{and} \quad \mathbf{b} = \begin{bmatrix} 0 \\ -2 \end{bmatrix}$$

and the corresponding homogeneous system (i.e. where we replace \mathbf{b} with $\mathbf{0}$.)

- (b) Find the solution sets of the system

$$A\mathbf{x} = \mathbf{b}$$

where

$$A = \begin{bmatrix} 1 & 2 & -1 \\ 2 & 4 & -2 \end{bmatrix}, \quad \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}, \quad \text{and} \quad \mathbf{b} = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$$

and the corresponding homogeneous system.

- (c) What are the similarities/differences between solutions of the nonhomogeneous system and its homogeneous counterpart?

As we saw in the above activity, there is a relationship between solutions of a nonhomogeneous and the corresponding homogeneous system. Let us formalize this relationship. If the general solution of a system involves free variables, we can represent the solutions in **parametric vector form** to have a better idea about the geometric representation of the solution set. Suppose the

solution is that x_3 is free, $x_2 = -2 + x_3$, and $x_1 = 2 - 3x_3$. In vector form, we can represent this general solution as

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 2 - 3x_3 \\ x_3 - 2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 2 \\ -2 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} -3 \\ 1 \\ 1 \end{bmatrix}. \quad (5.4)$$

From this representation, we see that the solution set is a line through the origin (formed by multiples of $\begin{bmatrix} -3 \\ 1 \\ 1 \end{bmatrix}$) shifted by the added vector $\begin{bmatrix} 2 \\ -2 \\ 0 \end{bmatrix}$. The solution to the homogeneous system on the other does not have the shift.

Algebraically, we see that every solution to the nonhomogeneous system $Ax = b$ can be written in the form $\mathbf{p} + \mathbf{v}_h$, where \mathbf{p} is a particular solution to $Ax = b$ and \mathbf{v}_h is a solution to the corresponding homogeneous system $Ax = \mathbf{0}$.

To understand why this always happens, we will verify the result algebraically for an arbitrary A and b . Assuming that \mathbf{p} is a particular solution to the nonhomogeneous system $Ax = b$, we need to show that:

- if \mathbf{v} is an arbitrary solution to the nonhomogeneous system, then $\mathbf{v} = \mathbf{p} + \mathbf{v}_h$, where \mathbf{v}_h is some solution to the homogeneous system $Ax = \mathbf{0}$, and
- if \mathbf{v}_h is an arbitrary solution to the homogeneous system, then $\mathbf{p} + \mathbf{v}_h$ is a solution to the nonhomogeneous system.

To verify the first condition, suppose that \mathbf{v} is a solution to the nonhomogeneous system $Ax = b$. Since we want $\mathbf{v} = \mathbf{p} + \mathbf{v}_h$, we need to verify that $\mathbf{v} - \mathbf{p}$ is a solution for the homogeneous system so that we can assign $v_h = \mathbf{v} - \mathbf{p}$. Note that

$$A(\mathbf{v} - \mathbf{p}) = A\mathbf{v} - A\mathbf{p} = \mathbf{b} - \mathbf{b} = \mathbf{0},$$

using the distributive property of matrix-vector product over vector addition. Hence \mathbf{v} is of the form $\mathbf{p} + \mathbf{v}_h$ with $\mathbf{v}_h = \mathbf{0}$.

To verify the second condition, consider a vector of the form $\mathbf{p} + \mathbf{v}_h$, where \mathbf{v}_h is a homogeneous solution. We have

$$A(\mathbf{p} + \mathbf{v}_h) = A\mathbf{p} + A\mathbf{v}_h = \mathbf{b} + \mathbf{0} = \mathbf{b},$$

and so $\mathbf{p} + \mathbf{v}_h$ is a solution to $Ax = b$.

Our work above proves the following theorem.

Theorem 5.4. *Suppose the equation $Ax = b$ is consistent for some b and \mathbf{p} is a solution. Then the solution set of $Ax = b$ consists of all vectors of the form $\mathbf{v} = \mathbf{p} + \mathbf{v}_h$ where \mathbf{v}_h is a solution to $Ax = \mathbf{0}$.*



The Geometry of Solutions to the Homogeneous System

There is a simple geometric interpretation to the solution set of the homogeneous system $A\mathbf{x} = \mathbf{0}$ based on the number of free variables that imposes a geometry on the solution set of the corresponding nonhomogeneous system $A\mathbf{x} = \mathbf{b}$ (when consistent) due to Theorem 5.4.

Activity 5.4. In this activity we consider geometric interpretations of the solution sets of homogeneous and nonhomogeneous systems.

- (a) Consider the system $A\mathbf{x} = \mathbf{b}$ where $A = \begin{bmatrix} 1 & -3 \\ -3 & 9 \\ -1 & 3 \end{bmatrix}$ and $\mathbf{b} = \begin{bmatrix} 2 \\ -6 \\ -2 \end{bmatrix}$. The general solution to this system has the form $\begin{bmatrix} 2 \\ 0 \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} 3 \\ 1 \\ 1 \end{bmatrix}$, where x_2 is any real number.

i. Let $\mathbf{v} = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$. What does the set of all vectors of the form $x_2\mathbf{v}$ look like geometrically? Draw a picture in \mathbb{R}^2 to illustrate. (Recall that we refer to all the vectors of the form $x_2\mathbf{v}$ simply as $\text{Span}\{\mathbf{v}\}$.)

ii. Let $\mathbf{p} = \begin{bmatrix} 2 \\ 0 \end{bmatrix}$. What effect does adding the vector \mathbf{p} to each vector in $\text{Span}\{\mathbf{v}\}$ have on the geometry of $\text{Span}\{\mathbf{v}\}$? Finally, what does this mean about the geometry of the solution set to the nonhomogeneous system $A\mathbf{x} = \mathbf{b}$?

- (b) Consider the system $A\mathbf{x} = \mathbf{b}$ where $A = \begin{bmatrix} 1 & 2 & -1 \\ 3 & 6 & -3 \end{bmatrix}$ and $\mathbf{b} = \begin{bmatrix} -2 \\ -6 \end{bmatrix}$. The general solution to this system has the form $\begin{bmatrix} -2 \\ 0 \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$, where x_2, x_3 are any real numbers.

i. Let $\mathbf{u} = \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix}$, $\mathbf{v} = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$. Use our results from Section 4 to determine the geometric shape of $\text{Span}\{\mathbf{u}, \mathbf{v}\}$, the set of all vectors of the form $x_2 \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$, where x_2, x_3 are any real numbers.

ii. Let $\mathbf{p} = \begin{bmatrix} -2 \\ 0 \\ 0 \end{bmatrix}$. What's the geometric effect of adding the vector \mathbf{p} to each vector in $\text{Span}\{\mathbf{u}, \mathbf{v}\}$? Finally, what does this mean about the geometry of the solution set to the nonhomogeneous system $A\mathbf{x} = \mathbf{b}$?

Our work in the above activity shows the geometric shape of the solution set of a consistent nonhomogeneous system is the same as the geometric shape of the solution set of the corresponding homogeneous system. The only difference between the two solution sets is that one is a shifted version of the other.

Examples

What follows are worked examples that use the concepts from this section.

Example 5.5. We now have several different ways to represent a system of linear equations. Rewrite the system in an equivalent form

$$\begin{aligned} 11x_1 + 4x_2 - 5x_3 - 2x_4 &= 63 \\ 15x_1 + 5x_2 + 2x_3 - 2x_4 &= 68 \\ 6x_1 + 2x_2 + x_3 - x_4 &= 26 \\ 9x_1 + 3x_2 + 2x_3 - x_4 &= 40. \end{aligned}$$

- (a) as an augmented matrix
- (b) as an equation involving a linear combination of vectors
- (c) using a matrix-vector product

Then solve the system.

Example Solution.

- (a) The augmented matrix for this system is

$$\left[\begin{array}{cccc|c} 11 & 4 & -5 & -2 & 63 \\ 15 & 5 & 2 & -2 & 68 \\ 6 & 2 & 1 & -1 & 26 \\ 9 & 3 & 2 & -1 & 40 \end{array} \right].$$

- (b) If we make vectors from the columns of the augmented matrix, we can write this system in vector form as

$$x_1 \begin{bmatrix} 11 \\ 15 \\ 6 \\ 9 \end{bmatrix} + x_2 \begin{bmatrix} 4 \\ 5 \\ 2 \\ 3 \end{bmatrix} + x_3 \begin{bmatrix} -5 \\ 2 \\ 1 \\ 2 \end{bmatrix} + x_4 \begin{bmatrix} -2 \\ -2 \\ -1 \\ -1 \end{bmatrix} = \begin{bmatrix} 63 \\ 68 \\ 26 \\ 40 \end{bmatrix}.$$

- (c) The coefficient matrix for this system is $\begin{bmatrix} 11 & 4 & -5 & -2 \\ 15 & 5 & 2 & -2 \\ 6 & 2 & 1 & -1 \\ 9 & 3 & 2 & -1 \end{bmatrix}$, and the matrix-vector form of the system is

$$\begin{bmatrix} 11 & 4 & -5 & -2 \\ 15 & 5 & 2 & -2 \\ 6 & 2 & 1 & -1 \\ 9 & 3 & 2 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 63 \\ 68 \\ 26 \\ 40 \end{bmatrix}.$$

Using technology, we find that the reduced row echelon form of the augmented matrix for this system is

$$\left[\begin{array}{cccc|c} 1 & 0 & 0 & 0 & 3 \\ 0 & 1 & 0 & 0 & 7 \\ 0 & 0 & 1 & 0 & -2 \\ 0 & 0 & 0 & 1 & 4 \end{array} \right].$$

So the solution to this system is $x_1 = 3$, $x_2 = 7$, $x_3 = -2$, and $x_4 = 4$.

Example 5.6. Consider the homogeneous system

$$\begin{aligned} x_1 + 8x_2 - x_3 &= 0 \\ x_1 - 7x_2 + 2x_3 &= 0 \\ 3x_1 + 4x_2 + x_3 &= 0. \end{aligned}$$

- (a) Find the general solution to this homogeneous system and express the system in parametric vector form.
- (b) Let $A = \begin{bmatrix} 1 & 8 & -1 \\ 1 & -7 & 2 \\ 3 & 4 & 1 \end{bmatrix}$, and let $\mathbf{b} = \begin{bmatrix} -6 \\ 9 \\ 2 \end{bmatrix}$. Show that $\begin{bmatrix} -1 \\ 0 \\ 5 \end{bmatrix}$ is a solution to the non-homogeneous system $A\mathbf{x} = \mathbf{b}$.
- (c) Use the results from part (a) and (b) to write the parametric vector form of the general solution to the non-homogeneous system $A\mathbf{x} = \mathbf{b}$. (Do this without directly solving the system $A\mathbf{x} = \mathbf{b}$.)
- (d) Describe what the general solution to the homogeneous system $A\mathbf{x} = \mathbf{0}$ and the general solution to the non-homogeneous system $A\mathbf{x} = \mathbf{b}$ look like geometrically.

Example Solution.

- (a) The augmented matrix of the homogeneous system is

$$\left[\begin{array}{ccc|c} 1 & 8 & -1 & 0 \\ 1 & -7 & 2 & 0 \\ 3 & 4 & 1 & 0 \end{array} \right],$$

and the reduced row echelon form of this augmented matrix is

$$\left[\begin{array}{ccc|c} 1 & 0 & \frac{3}{5} & 0 \\ 1 & -\frac{1}{5} & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array} \right].$$

Since there is no corresponding equation of the form $0 = b$ for a nonzero constant b , this system is consistent. The third column contains no pivot, so the variable x_3 is free, $x_2 =$

$\frac{1}{5}x_3$ and $x_1 = -\frac{3}{5}x_3$. In parametric vector form the general solution to the homogeneous system is

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -\frac{3}{5}x_3 \\ \frac{1}{5}x_3 \\ x_3 \end{bmatrix} = x_3 \begin{bmatrix} -\frac{3}{5} \\ \frac{1}{5} \\ 1 \end{bmatrix}.$$

(b) Since

$$\begin{aligned} A \begin{bmatrix} -1 \\ 0 \\ 5 \end{bmatrix} &= (-1) \begin{bmatrix} 1 \\ 1 \\ 3 \end{bmatrix} + (0) \begin{bmatrix} 8 \\ -7 \\ 4 \end{bmatrix} + (5) \begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix} \\ &= \begin{bmatrix} -1 - 5 \\ -1 + 10 \\ -3 + 5 \end{bmatrix} = \begin{bmatrix} -6 \\ 9 \\ 2 \end{bmatrix}, \end{aligned}$$

we conclude that $\begin{bmatrix} -1 \\ 0 \\ 5 \end{bmatrix}$ is a solution to the non-homogeneous system $A\mathbf{x} = \mathbf{b}$.

(c) We know that every solution to the non-homogeneous system $A\mathbf{x} = \mathbf{b}$ has the form of the general solution to the homogeneous system plus a particular solution to the non-homogeneous system. Combining the results of (a) and (b) we see that the general solution to the non-homogeneous system $A\mathbf{x} = \mathbf{b}$ is

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -1 \\ 0 \\ 5 \end{bmatrix} + x_3 \begin{bmatrix} -\frac{3}{5} \\ \frac{1}{5} \\ 1 \end{bmatrix},$$

where x_3 can be any real number.

(d) The solution to the homogeneous system $A\mathbf{x} = \mathbf{0}$ is the span of the vector $\begin{bmatrix} -\frac{3}{5} \\ \frac{1}{5} \\ 1 \end{bmatrix}$.

Geometrically, this set of points is a line through the origin and the point $(-3, 1, 5)$ in \mathbb{R}^3 . The solution to the non-homogeneous system $A\mathbf{x} = \mathbf{b}$ is the translation of the line

through the origin and $(-3, 1, 5)$ by the vector $\begin{bmatrix} -1 \\ 0 \\ 5 \end{bmatrix}$. In other words, the solution to

the non-homogeneous system $A\mathbf{x} = \mathbf{b}$ is the line in \mathbb{R}^3 through the points $(-1, 0, 5)$ and $(-4, 1, 10)$.

Summary

- If $A = [\mathbf{c}_1 \ \mathbf{c}_2 \ \cdots \ \mathbf{c}_n]$ is an $m \times n$ matrix with columns $\mathbf{c}_1, \mathbf{c}_2, \dots, \mathbf{c}_n$, and if $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$ is a vector in \mathbb{R}^n , then the matrix-vector product $A\mathbf{x}$ is defined to be the linear combination of the columns of A with corresponding weights from \mathbf{x} – that is

$$A\mathbf{x} = x_1\mathbf{c}_1 + x_2\mathbf{c}_2 + \cdots + x_n\mathbf{c}_n.$$

- A linear system

$$a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1$$

$$a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2$$

⋮

$$a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m$$

can be written in matrix form as

$$A\mathbf{x} = \mathbf{b},$$

where

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \ddots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}, \quad \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, \quad \text{and} \quad \mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}.$$

- The matrix equation $A\mathbf{x} = \mathbf{b}$ has a solution if and only if \mathbf{b} is a linear combination of the columns of A .
- The system $A\mathbf{x} = \mathbf{b}$ is consistent for every vector \mathbf{b} if every row of A contains a pivot.
- A homogeneous system is a system of the form $A\mathbf{x} = \mathbf{0}$ for some $m \times n$ matrix A . Since the zero vector in \mathbb{R}^n satisfies $A\mathbf{x} = \mathbf{0}$, a homogeneous system is always consistent.
- A homogeneous system can have one or infinitely many different solutions. The homogeneous system $A\mathbf{x} = \mathbf{0}$ has exactly one solution if and only if each column of A is a pivot column.
- The solutions to the consistent nonhomogeneous system $A\mathbf{x} = \mathbf{b}$ have the form $\mathbf{p} + \mathbf{v}_h$, where \mathbf{p} is a particular solution to the nonhomogeneous system $A\mathbf{x} = \mathbf{b}$ and \mathbf{v}_h is a solution to the homogeneous system $A\mathbf{x} = \mathbf{0}$. In other words, the solution space to a consistent nonhomogeneous system $A\mathbf{x} = \mathbf{b}$ is a translation of the solution space of the homogeneous system $A\mathbf{x} = \mathbf{0}$ by a particular solution to the nonhomogeneous system.

Finally, we argued an important theorem.



Theorem 5.7. Let A be an $m \times n$ matrix. The following statements are equivalent.

- (1) The matrix equation $A\mathbf{x} = \mathbf{b}$ has a solution for every vector \mathbf{b} in \mathbb{R}^m .
- (2) Every vector \mathbf{b} in \mathbb{R}^m can be written as a linear combination of the columns of A .
- (3) The span of the columns of A is \mathbb{R}^m .
- (4) The matrix A has a pivot position in each row.

We will continue to add to this theorem, so it is a good idea for you to begin now to remember the equivalent conditions of this theorem.

Exercises

- (1) Write the system

$$\begin{aligned}x_1 + 2x_2 + 2x_3 + x_4 &= -1 \\4x_1 - 8x_2 + 3x_3 - 9x_4 &= 2 \\x_1 + 6x_2 - 4x_3 + 12x_4 &= -1\end{aligned}$$

in matrix-vector form. Explicitly identify the coefficient matrix and the vector of constants.

- (2) Write the linear combination

$$x_1 \begin{bmatrix} 1 \\ 5 \end{bmatrix} + x_2 \begin{bmatrix} -3 \\ 10 \end{bmatrix} + x_3 \begin{bmatrix} 2 \\ 2 \end{bmatrix}$$

as a matrix-vector product.

- (3) Represent the following matrix-vector equation as a linear system and find its solution.

$$\begin{bmatrix} 2 & 3 & 4 \\ 1 & -2 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 4 \\ -6 \end{bmatrix}$$

- (4) Represent the following matrix-vector equation as a linear system and find its solution.

$$\begin{bmatrix} 1 & -2 & -1 \\ 2 & 2 & -2 \\ 3 & 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 \\ -4 \\ 8 \end{bmatrix}$$

- (5) Another way of defining the matrix-vector product uses the concept of the *scalar product* of

vectors.² Given a $1 \times n$ matrix $\mathbf{u} = [u_1 \ u_2 \ \dots \ u_n]$ ³ and an $n \times 1$ vector $\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$, we define the scalar product $\mathbf{u} \cdot \mathbf{v}$ as

$$\mathbf{u} \cdot \mathbf{v} = u_1 v_1 + u_2 v_2 + u_3 v_3 + \dots + u_n v_n.$$

We then define the matrix-vector product $A\mathbf{x}$ as the vector whose entries are the scalar products of the rows of A with \mathbf{x} . As an example, if $A = \begin{bmatrix} 2 & 3 & 4 \\ 1 & -2 & 3 \end{bmatrix}$ and $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$, then

$$A\mathbf{x} = \begin{bmatrix} 2x_1 + 3x_2 + 4x_3 \\ x_1 + (-2)x_2 + 3x_3 \end{bmatrix}.$$

Calculate the matrix-vector product $A\mathbf{x}$ where $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ and $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ using both methods of finding the matrix-vector product to show that the two definitions are equivalent for size 2×2 matrices.

- (6) Find the value of a such that

$$\begin{bmatrix} 1 & 2 & 2 \\ 1 & -1 & 3 \\ 1 & 2 & 4 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \\ a \end{bmatrix} = \begin{bmatrix} * \\ -5 \\ * \end{bmatrix}$$

where $*$'s represent unknown values.

- (7) Suppose we have

$$\begin{bmatrix} 1 & 2 & 1 & 2 \\ -1 & 2 & 3 & 1 \\ 2 & 3 & 1 & a \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ -2 \\ 3 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$$

where b_i 's represent unknown values.

(a) In order to find the value of a , which of the b_i 's do we need to know? Why?

(b) Suppose the b_i 's that we need to know is(are) equal to 9. What is the value of a ?

- (8) Suppose we are given

$$A\mathbf{u} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \text{ and } A\mathbf{v} = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$$

for an unknown A and two unknown vectors \mathbf{u}, \mathbf{v} in \mathbb{R}^3 . Using matrix-vector product properties, evaluate $A\mathbf{w}$ where $\mathbf{w} = 2\mathbf{u} - 3\mathbf{v}$.

²Note that some authors refer to the scalar product as the *dot product*.

³We can identify a $1 \times n$ matrix $\mathbf{u} = [u_1 \ u_2 \ \dots \ u_n]$ with the $n \times 1$ vector $\mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix}$, so we often refer to $[u_1 \ u_2 \ \dots \ u_n]$ as a vector.

(9) Suppose we are given

$$A \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \text{ and } A \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix} = \begin{bmatrix} 0 \\ 2 \\ 2 \end{bmatrix}.$$

After expressing $\begin{bmatrix} -1 \\ 6 \\ -5 \end{bmatrix}$ as a linear combination of $\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ and $\begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix}$, use the matrix-vector product properties to determine $A \begin{bmatrix} -1 \\ 6 \\ -5 \end{bmatrix}$.

(10) (a) The non-homogeneous system (with unknown constants a and b)

$$\begin{aligned} x + y - z &= 2 \\ 2x + ay + bz &= 4 \end{aligned}$$

has a solution which lies on the x -axis (i.e. $y = z = 0$). Find this solution.

(b) If the corresponding homogeneous system

$$\begin{aligned} x + y - z &= 0 \\ 2x + ay + bz &= 0 \end{aligned}$$

has its general solution expressed in parametric vector form as $z \cdot \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$, find the general solution for the non-homogeneous system using your answer to part (a).

(c) Find the conditions on a and b that make the system from (a) have the general solution you found in (b).

(11) Find the general solution to the non-homogeneous system

$$\begin{aligned} x - 2y + z &= 3 \\ -2x + 4y - 2z &= -6. \end{aligned}$$

Using the parametric vector form of the solutions, determine what the solution set to this non-homogeneous system looks like geometrically. Be as specific as possible. (Include information such as whether the solution set is a point, a line, or a plane, etc.; whether the solution set passes through the origin or is shifted from the origin in a specific direction by a specific number of units; and how the solution is related to the corresponding homogeneous system.)

- (12) Come up with an example of a 3×3 matrix A for which the solution set of $A\mathbf{x} = \mathbf{0}$ is a line, and a 3×3 matrix A for which the solution set of $A\mathbf{x} = \mathbf{0}$ is a plane.
- (13) Suppose we have three vectors $\mathbf{v}_1, \mathbf{v}_2$ and \mathbf{v}_3 satisfying $\mathbf{v}_3 = 2\mathbf{v}_1 - \mathbf{v}_2$. Let A be the matrix with vectors $\mathbf{v}_1, \mathbf{v}_2$ and \mathbf{v}_3 as the columns in that order. Find a non-zero \mathbf{x} such that $A\mathbf{x} = \mathbf{0}$ using this information.



(14) Label each of the following statements as True or False. Provide justification for your response.

- (a) **True/False** If the system $Ax = \mathbf{0}$ has infinitely many solutions, then so does the system $Ax = \mathbf{b}$ for **any** right-hand-side \mathbf{b} .
- (b) **True/False** If \mathbf{x}_1 is a solution for $Ax = \mathbf{b}_1$ and \mathbf{x}_2 is a solution for $Ax = \mathbf{b}_2$, then $\mathbf{x}_1 + \mathbf{x}_2$ is a solution for $Ax = \mathbf{b}_1 + \mathbf{b}_2$.
- (c) **True/False** If an $m \times n$ matrix A has a pivot in every row, then the equation $Ax = \mathbf{b}$ has a unique solution for every \mathbf{b} .
- (d) **True/False** If an $m \times n$ matrix A has a pivot in every row, then the equation $Ax = \mathbf{b}$ has a solution for every \mathbf{b} .
- (e) **True/False** If A and B are row equivalent matrices and the columns of A span \mathbb{R}^m , then so do the columns of B .
- (f) **True/False** All homogeneous systems have either a unique solution or infinitely many solutions.
- (g) **True/False** If a linear system is not homogeneous, then the solution set does not include the origin.
- (h) **True/False** If a solution set of a linear system does not include the origin, the system is not homogeneous.
- (i) **True/False** If the system $Ax = \mathbf{b}$ has a unique solution for some \mathbf{b} , then the homogeneous system has only the trivial solution.
- (j) **True/False** If A is a 3×4 matrix, then the homogeneous equation $Ax = \mathbf{0}$ has non-trivial solutions.
- (k) **True/False** If A is a 3×2 matrix, then the homogeneous equation $Ax = \mathbf{0}$ has non-trivial solutions.

Project: Input-Output Models

There are two basic types of input-output models: closed and open. The closed model assumes that all goods produced are consumed within the economy – no trading takes place with outside entities. In the open model, goods produced within the economy can be traded outside the economy.

To work with a closed model, we use an example (from *Input-Output Economics* by Wassily Leontief). Assume a simple three-sector economy consisting of agriculture (growing wheat), manufacturing (producing cloth), and households (supplying labor). Each sector of the economy relies on goods from the other sectors to operate (e.g., people must eat to work and need to be clothed). To model the interactions between the sectors, we consider how many units of product is needed as input from one sector to another to produce one unit of product in the second sector. For example, assume the following:

- to produce one unit (say dollars worth) of agricultural goods requires 25% of a unit of agricultural output, 28% of a unit of manufacturing output, and 27% of a unit of household output;



- to produce one unit of manufactured goods requires 20% of a unit of agricultural output, 60% of a unit of manufacturing output, and 60% of a unit of household output;
- to produce one unit of household goods requires 55% of a unit of agricultural output, 12% of a unit of manufacturing output, and 13% of a unit of household output.

These assumptions are summarized in Table 5.1.

into \ from	Agriculture	Manufacture	Households
Agriculture	0.25	0.28	0.27
Manufacture	0.20	0.60	0.60
Households	0.55	0.12	0.13

Table 5.1: Summary of simple three sector economy.

This model is said to be *closed* because all good produced are used up within the economy. If there are goods that are not used within the economy the model is said to be *open*. Open models will be examined later.

The economist's goal is to determine what level of production in each section meets the following requirements:

- the production from each sector meets the needs of all of the sectors and
- there is no overproduction.

Project Activity 5.1. We can use techniques from linear algebra to determine the levels of production that precisely meet the two goals of the economist.

- (a) Suppose that the agricultural output is x_1 units, the manufacturing output is x_2 units, and the household output is x_3 units. We represent this data as a *production vector* $\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$. To produce a unit of agriculture requires 0.25 units of agriculture, 0.28 units of manufacturing, and 0.27 units of household. If x_1 units of agriculture, x_2 units of manufacturing, and x_3 units of household products are produced, then agriculture can produce

$$0.25x_1 + 0.28x_2 + 0.27x_3$$

units. In order to meet the needs of agriculture and for there to be no overproduction, we must then have

$$0.25x_1 + 0.28x_2 + 0.27x_3 = x_1.$$

Write similar equations for the manufacturing and household sectors of the economy.

- (b) Find the augmented matrix for the system of linear equations that represent production of the three sectors from part (a), and then solve the system to find the production levels that meet the economist's two goals.



- (c) Suppose the production level of the household sector is 200 million units (dollars). Find the production levels of the agricultural and manufacturing sectors that meet the economist's two goals.

In general, a matrix derived from a table like Table 5.1 is called a *consumption* matrix, which we will denote as C . (In the example discussed here $C = \begin{bmatrix} 0.25 & 0.28 & 0.27 \\ 0.20 & 0.60 & 0.60 \\ 0.55 & 0.12 & 0.13 \end{bmatrix}$.) A consumption matrix $C = [c_{ij}]$, where c_{ij} represents the proportion of the output of sector j that is consumed by sector i , satisfies two important properties.

- Since no sector can consume a negative amount or an amount that exceeds the output of another sector, we must have $0 \leq c_{ij} \leq 1$ for all i and j .
- If there are n sectors in the economy, the fact that all output is consumed within the economy implies that $c_{1j} + c_{2j} + \cdots + c_{nj} = 1$. In other words, the column sums of C are all 1.

In our example, if we let $x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$, then we can write the equations that guarantee that the production levels satisfy the two economists' goal in matrix form as

$$\mathbf{x} = C\mathbf{x}. \quad (5.5)$$

Now we can rephrase the question to be answered as which production vectors \mathbf{x} satisfy equation (5.5). When $C\mathbf{x} = \mathbf{x}$, then the system is in equilibrium, that is output exactly meets needs. Any solution \mathbf{x} that satisfies (5.5) is called a *steady state* solution.

Project Activity 5.2. Is there a steady state solution for the closed system of Agriculture, Manufacturing, and Households? If so, find the general steady state solution. If no, explain why.

So far, we considered the case where the economic system was *closed*. This means that the industries that were part of the system sold products only to each other. However, if we want to represent the demand from other countries, from households, capital building, etc., we need an *open model*. In an article in the *Scientific American* Leontief organized the 1958 American economy into 81 sectors. The production of each of these sectors relied on production from all of the sectors. Here we present a small sample from Leontief's 81 sectors, using Petroleum, Textiles, Transportation, and Chemicals as our sectors of the economy. Leontief's model assumed that the production of 1 unit of output of

- petroleum requires 0.1 unit of petroleum, 0.2 units of transportation, and 0.4 units of chemicals;
- textiles requires 0.4 units of petroleum, 0.1 unit of textiles, 0.15 units of transportation, and 0.3 units of chemicals;
- transportation requires 0.6 units of petroleum, 0.1 unit of transportation, and 0.25 units of chemicals;



- chemicals requires 0.2 units of petroleum, 0.1 unit of textiles, 0.3 units of transportation, and 0.2 units of chemicals.

A summary of this information is in Table 5.2. Assume the units are measured in dollars.

into \ from	Petroleum	Textiles	Transportation	Chemicals
Petroleum	0.10	0.00	0.20	0.40
Textiles	0.40	0.10	0.15	0.30
Transportation	0.60	0.00	0.10	0.25
Chemicals	0.20	0.10	0.30	0.20

Table 5.2: Summary of four sector economy.

In the open model, there is another part of the economy, called the *open sector*, that does not produce goods or services but only consumes them. If this sector (think end consumers, for example) demands/consumes d_1 units of Petroleum, d_2 units of Textiles, d_3 units of Transportation,

and d_4 units of Chemicals, we put this into a *final demand vector* $\mathbf{d} = \begin{bmatrix} d_1 \\ d_2 \\ d_3 \\ d_4 \end{bmatrix}$.

An economist would want to find the production level where the demand from the good/service producing sectors of the economy plus the final demand from the open sector exactly matches the output in each of the sectors. Let x_1 represent the number of units of petroleum output, x_2 the number of units of textiles output, x_3 the number of units of transportation output, and x_4 the number of units of chemical output during any time period. Then the production vector is

$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}$. So an economist wants to find the production vectors \mathbf{x} such that

$$\begin{aligned} 0.10x_1 + 0.20x_3 + 0.40x_4 + d_1 &= x_1 \\ 0.40x_1 + 0.10x_2 + 0.15x_3 + 0.30x_4 + d_2 &= x_2 \\ 0.60x_1 + 0.10x_3 + 0.25x_4 + d_3 &= x_3 \\ 0.20x_1 + 0.10x_2 + 0.30x_3 + 0.20x_4 + d_4 &= x_4, \end{aligned}$$

where $\mathbf{d} = \begin{bmatrix} d_1 \\ d_2 \\ d_3 \\ d_4 \end{bmatrix}$ is the demand vector from the open market. The matrix

$$E = \begin{bmatrix} 0.10 & 0.00 & 0.20 & 0.40 \\ 0.40 & 0.10 & 0.15 & 0.30 \\ 0.60 & 0.00 & 0.10 & 0.25 \\ 0.20 & 0.10 & 0.30 & 0.20 \end{bmatrix}$$



derived from Table 5.2, is called the *exchange* matrix.

Project Activity 5.3.

- (a) Suppose the final demand vector in our four sector economy is $\begin{bmatrix} 500 \\ 200 \\ 400 \\ 100 \end{bmatrix}$. Find the production levels that satisfy our system.
- (b) Does this economy defined by the exchange matrix E have production levels that exactly meet internal and external demands regardless of the external demands? That is, does the system of equations

$$\begin{aligned} 0.10x_1 + 0.20x_3 + 0.40x_4 + d_1 &= x_1 \\ 0.40x_1 + 0.10x_2 + 0.15x_3 + 0.30x_4 + d_2 &= x_2 \\ 0.60x_1 + 0.10x_3 + 0.25x_4 + d_3 &= x_3 \\ 0.20x_1 + 0.10x_2 + 0.30x_3 + 0.20x_4 + d_4 &= x_4 \end{aligned}$$

have a solution regardless of the values of d_1 , d_2 , d_3 , and d_4 ? Explain.

Section 6

Linear Dependence and Independence

Focus Questions

By the end of this section, you should be able to give precise and thorough answers to the questions listed below. You may want to keep these questions in mind to focus your thoughts as you complete the section.

- What are two ways to describe what it means for a set of vectors in \mathbb{R}^n to be linearly independent?
- What are two ways to describe what it means for a set of vectors in \mathbb{R}^n to be linearly dependent?
- If S is a set of vectors, what do we mean by a basis for $\text{Span } S$?
- Given a nonzero set S of vectors, how can we find a linearly independent subset of S that has the same span as S ?
- How do we recognize if the columns of a matrix A are linearly independent?
- How can we use a matrix to determine if a set $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k\}$ of vectors is linearly independent?
- How can we use a matrix to find a minimal spanning set for a set $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \dots, \mathbf{v}_k\}$ of vectors in \mathbb{R}^n ?

Application: Bézier Curves

Bézier curves are simple curves that were first developed in 1959 by French mathematician Paul de Casteljau, who was working at the French automaker Citroën. The curves were made public in 1962 by Pierre Bézier who used them in his work designing automobiles at the French car maker Renault. In addition to automobile design, Bézier curves have many other uses. Two of the most common applications of Bézier curves are font design and drawing tools. As an example, the letter

“S” in Palatino font is shown using Bézier curves in Figure 6.1. If you’ve used Adobe Illustrator, Photoshop, Macromedia Freehand, Fontographer, or any other of a number of drawing programs, then you’ve used Bézier curves. At the end of this section we will see how Bézier curves can be defined using linearly independent vectors and linear combinations of vectors.

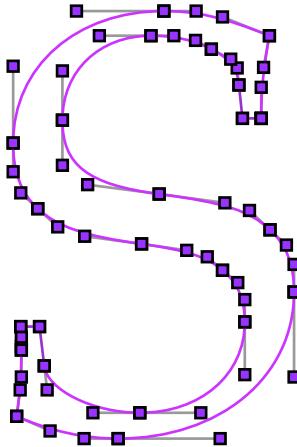


Figure 6.1: A letter *S*.

Introduction

In Section 4 we saw how to represent water-benzene-acetic acid chemical solutions with vectors, where the components represent the water, benzene and acid percentages. We then considered a problem of determining if a given chemical solution could be made by mixing other chemical solutions. Suppose we now have three different water-benzene-acetic acid chemical solutions, one with 40% water, 50% benzene and 10% acetic acid, the second with 52% water, 42% benzene and 6% acid, and a third with 46% water, 46% benzene and 8% acid. We represent the first chemical

solution with the vector $\mathbf{v}_1 = \begin{bmatrix} 40 \\ 50 \\ 10 \end{bmatrix}$, the second with the vector $\mathbf{v}_2 = \begin{bmatrix} 52 \\ 42 \\ 6 \end{bmatrix}$, and the third with the vector $\mathbf{v}_3 = \begin{bmatrix} 46 \\ 46 \\ 8 \end{bmatrix}$. By combining these three chemical solutions we can make a chemical

solution with 43% water, 48% benzene and 9% acid as follows

$$\frac{7}{12}\mathbf{v}_1 + \frac{1}{12}\mathbf{v}_2 + \frac{1}{3}\mathbf{v}_3 = \begin{bmatrix} 43 \\ 48 \\ 9 \end{bmatrix}.$$

However, if we had noticed that the third chemical solution can actually be made from the first two, that is,

$$\frac{1}{2}\mathbf{v}_1 + \frac{1}{2}\mathbf{v}_2 = \mathbf{v}_3,$$

we might have realized that we don't need the third chemical solution to make the 43% water, 48% benzene and 9% acid chemical solution. In fact,

$$\frac{3}{4}\mathbf{v}_1 + \frac{1}{4}\mathbf{v}_2 = \begin{bmatrix} 43 \\ 48 \\ 9 \end{bmatrix}.$$

(See Exercise 5 of Section 4.) Using the third chemical solution (represented by \mathbf{v}_3) uses more information than we actually need to make the desired 43% water, 48% benzene and 9% acid chemical solution because the vector \mathbf{v}_3 is redundant – all of the material we need to make \mathbf{v}_3 is contained in \mathbf{v}_1 and \mathbf{v}_2 . This is the basic idea behind linear independence – representing information in the most efficient way.

Information is often contained in and conveyed through vectors – especially linear combinations of vectors. In this section we will investigate the concepts of linear dependence and independence of a set of vectors. Our goal is to be able to efficiently determine when a given set of vectors forms a *minimal spanning set*. A minimal spanning set is a spanning set that contains the smallest number of vectors to obtain all of the vectors in the span. An important aspect of a minimal spanning set is that every vector in the span can be written in one and only one way as a linear combination of the vectors in the minimal spanning set. This will allow us to define the important notion of the dimension of a vector space.

Review of useful information: Recall that a linear combination of vectors $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k$ in \mathbb{R}^n is a sum of scalar multiples of $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k$. That is, a linear combination of the vectors $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k$ is a vector of the form

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_k\mathbf{v}_k,$$

where c_1, c_2, \dots, c_k are scalars.

Recall also that the collection of all linear combinations of a set $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k\}$ of vectors in \mathbb{R}^n is called the span of the set of vectors. That is, the span $\text{Span}\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k\}$ of the set $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k$ of vectors in \mathbb{R}^n is the set

$$\{c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_k\mathbf{v}_k : \text{ where } c_1, c_2, \dots, c_k \text{ are scalars}\}.$$

For example, a linear combination of vectors $\mathbf{v}_1 = \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}$ and $\mathbf{v}_2 = \begin{bmatrix} 0 \\ -2 \\ 1 \end{bmatrix}$ is $2\mathbf{v}_1 - 3\mathbf{v}_2 = \begin{bmatrix} 2 \\ 8 \\ 1 \end{bmatrix}$. All linear combinations of these two vectors can be expressed as the collection of vectors of the form $\begin{bmatrix} c_1 \\ c_1 - 2c_2 \\ 2c_1 + c_2 \end{bmatrix}$ where c_1, c_2 are scalars. Suppose we want to determine whether $\mathbf{w} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$ is in the span, in other words if \mathbf{w} is a linear combination of $\mathbf{v}_1, \mathbf{v}_2$. This means we are



looking for c_1, c_2 such that

$$\begin{bmatrix} c_1 \\ c_1 - 2c_2 \\ 2c_1 + c_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}.$$

we solve for the system represented with the augmented matrix

$$\left[\begin{array}{cc|c} 1 & 0 & 1 \\ 1 & -2 & 2 \\ 2 & 1 & 3 \end{array} \right].$$

By reducing this matrix, we find that there are no solutions of the system, which implies that \mathbf{w} is not a linear combination of $\mathbf{v}_1, \mathbf{v}_2$. Note that we can use any names we please for the scalars, say x_1, x_2 , if we prefer.

Preview Activity 6.1. Let $\mathbf{v}_1 = \begin{bmatrix} 2 \\ 1 \\ -3 \end{bmatrix}$, $\mathbf{v}_2 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$, and $\mathbf{v}_3 = \begin{bmatrix} 1 \\ -1 \\ -6 \end{bmatrix}$, and let $\mathbf{b} = \begin{bmatrix} 0 \\ 1 \\ 3 \end{bmatrix}$.

If \mathbf{b} is in $\text{Span}\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$, we are interested in the most efficient way to represent \mathbf{b} as a linear combination of $\mathbf{v}_1, \mathbf{v}_2$, and \mathbf{v}_3 .

- (1) The vector \mathbf{b} is in $\text{Span}\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ if there exist x_1, x_2 , and x_3 so that

$$x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + x_3\mathbf{v}_3 = \mathbf{b}.$$

(Recall that we can use any letters we want for the scalars. They are simply unknown scalars we want to solve for.)

(a) Explain why \mathbf{b} is in $\text{Span}\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$. (Hint: What is the matrix we need to reduce?)

(b) Write \mathbf{b} as a linear combination of $\mathbf{v}_1, \mathbf{v}_2$, and \mathbf{v}_3 . In how many ways can \mathbf{b} be written as a linear combination of the vectors $\mathbf{v}_1, \mathbf{v}_2$, and \mathbf{v}_3 ? Explain.

- (2) In problem 1 we saw that the vector \mathbf{b} could be written in infinitely many different ways as linear combinations of $\mathbf{v}_1, \mathbf{v}_2$, and \mathbf{v}_3 . We now ask the question if we really need all of the vectors $\mathbf{v}_1, \mathbf{v}_2$, and \mathbf{v}_3 to make \mathbf{b} as a linear combination in a unique way.

(a) Can the vector \mathbf{b} be written as a linear combination of the vectors \mathbf{v}_1 and \mathbf{v}_2 ? If not, why not? If so, in how many ways can \mathbf{b} be written as a linear combination of \mathbf{v}_1 and \mathbf{v}_2 ? Explain.

(b) If possible, write \mathbf{b} as a linear combination of \mathbf{v}_1 and \mathbf{v}_2 .

- (3) In problem 1 we saw that \mathbf{b} could be written in infinitely many different ways as a linear combination of the vectors $\mathbf{v}_1, \mathbf{v}_2$, and \mathbf{v}_3 . However, the vector \mathbf{b} could only be written in one way as a linear combination of \mathbf{v}_1 and \mathbf{v}_2 . So \mathbf{b} is in $\text{Span}\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ and \mathbf{b} is also in $\text{Span}\{\mathbf{v}_1, \mathbf{v}_2\}$. This raises a question – is *any* vector in $\text{Span}\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ also in $\text{Span}\{\mathbf{v}_1, \mathbf{v}_2\}$? If so, then the vector \mathbf{v}_3 is redundant in terms of forming the span of $\mathbf{v}_1, \mathbf{v}_2$, and \mathbf{v}_3 . For the sake of efficiency, we want to recognize and eliminate this redundancy.

(a) Can \mathbf{v}_3 be written as a linear combination of the vectors \mathbf{v}_1 and \mathbf{v}_2 ? If not, why not? If so, write \mathbf{v}_3 as a linear combination of \mathbf{v}_1 and \mathbf{v}_2 .

(b) Use the result of part (a) to decide if *any* vector in $\text{Span}\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is also in $\text{Span}\{\mathbf{v}_1, \mathbf{v}_2\}$.



Linear Independence

In this section we will investigate the concepts of linear dependence and independence of a set of vectors. Our goal is to be able to efficiently determine when a given set of vectors forms a *minimal spanning set*. This will involve the concepts of span and linear independence. Minimal spanning sets are important in that they provide the most efficient way to represent vectors in a space, and will later allow us to define the dimension of a vector space.

In Preview Activity 6.1 we considered the case where we had a set $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ of three vectors, and the vector \mathbf{v}_3 was in the span of $\{\mathbf{v}_1, \mathbf{v}_2\}$. So the vector \mathbf{v}_3 did not add anything to the span of $\{\mathbf{v}_1, \mathbf{v}_2\}$. In other words, the set $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ was larger than it needed to be in order to generate the vectors in its span – that is, $\text{Span}\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\} = \text{Span}\{\mathbf{v}_1, \mathbf{v}_2\}$. However, neither of the vectors in the set $\{\mathbf{v}_1, \mathbf{v}_2\}$ could be removed without changing its span. In this case, the set $\{\mathbf{v}_1, \mathbf{v}_2\}$ is what we will call a *minimal spanning set* or a *basis* for $\text{Span } S$. There are two important properties that make $\{\mathbf{v}_1, \mathbf{v}_2\}$ a basis for $\text{Span } S$. The first is that every vector in $\text{Span } S$ can be written as linear combinations of \mathbf{v}_1 and \mathbf{v}_2 (we also use the terminology that the vectors \mathbf{v}_1 and \mathbf{v}_2 span $\text{Span } S$), and the second is that every vector in $\text{Span } S$ can be written in exactly one way as a linear combination of \mathbf{v}_1 and \mathbf{v}_2 . This second property is the property of linear independence, and it is the property that makes the spanning set *minimal*.

To make a spanning set minimal, we want to be able to write every vector in the span in a unique way in terms of the spanning vectors. Notice that the zero vector can always be written as a linear combination of any set of vectors using 0 for all of the weights. So to have a *minimal* or *linearly independent* spanning set, that is, to have a unique representation for each vector in the span, it will need to be the case that the *only* way we can write the zero vector as a linear combination of a set of vectors is if all of the weights are 0. This leads us to the definition of a linearly independent set of vectors.

Definition 6.1. A set $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k\}$ of vectors in \mathbb{R}^n is **linearly independent** if the vector equation

$$x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + \cdots + x_k\mathbf{v}_k = \mathbf{0}$$

for scalars x_1, x_2, \dots, x_k has only the trivial solution

$$x_1 = x_2 = x_3 = \cdots = x_k = 0.$$

If a set of vectors is not linearly independent, then the set is **linearly dependent**.

Alternatively, we say that the vectors $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k$ are linearly independent (or dependent) if the set $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k\}$ is linearly independent (or dependent).

Note that the definition tells us that a set $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k\}$ of vectors in \mathbb{R}^n is linearly dependent if there are scalars x_1, x_2, \dots, x_n , not all of which are 0 so that

$$x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + \cdots + x_k\mathbf{v}_k = \mathbf{0}.$$

Activity 6.1. Which of the following sets in \mathbb{R}^2 or \mathbb{R}^3 is linearly independent and which is linearly dependent? Why? For the linearly dependent sets, write one of the vectors as a linear combination of the others, if possible.



$$(a) S_1 = \left\{ \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} -2 \\ 8 \\ 1 \end{bmatrix}, \begin{bmatrix} -4 \\ 8 \\ 0 \end{bmatrix} \right\}.$$

(b) $S_2 = \left\{ \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 2 \\ 3 \end{bmatrix} \right\}$. (Hint: What relationship must exist between two vectors if they are linearly dependent?)

(c) The vectors \mathbf{u} , \mathbf{v} , and \mathbf{w} as shown in Figure 6.2.

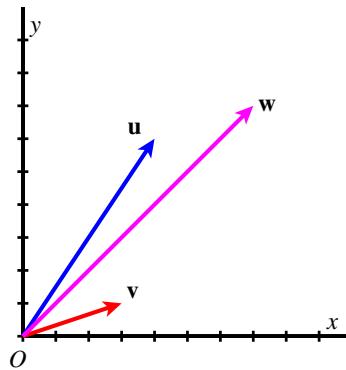


Figure 6.2: Vectors \mathbf{u} , \mathbf{v} , and \mathbf{w} .

Activity 6.1 (a) and (c) illustrate how we can write one of the vectors in a linearly dependent set as a linear combination of the others. This would allow us to write at least one of the vectors in the span of the set in more than one way as a linear combination of vectors in this set. We prove this result in general in the following theorem.

Theorem 6.2. *A set $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k\}$ of vectors in \mathbb{R}^n is linearly dependent if and only if at least one of the vectors in the set can be written as a linear combination of the remaining vectors in the set.*

The next activity is intended to help set the stage for the proof of Theorem 6.2.

Activity 6.2. The statement of Theorem 6.2 is a bi-conditional statement (an if and only if statement). To prove this statement about the set S we need to show two things about S . One: we must demonstrate that if S is a linearly dependent set, then at least one vector in S is a linear combination of the other vectors (this is the “only if” part of the biconditional statement) and Two: if at least one vector in S is a linear combination of the others, then S is linearly dependent (this is the “if” part of the biconditional statement). We illustrate the main idea of the proof using a three vector set $S = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$.

- (a) First let us assume that S is a linearly dependent set and show that at least one vector in S is a linear combination of the other vectors. Since S is linearly dependent we can write the zero vector as a linear combination of \mathbf{v}_1 , \mathbf{v}_2 , and \mathbf{v}_3 with at least one nonzero weight. For example, suppose

$$2\mathbf{v}_1 + 3\mathbf{v}_2 + 4\mathbf{v}_3 = \mathbf{0}. \quad (6.1)$$

Solve Equation (6.1) for the vector \mathbf{v}_2 to show that \mathbf{v}_2 can be written as a linear combination of \mathbf{v}_1 and \mathbf{v}_3 . Conclude that \mathbf{v}_2 is a linear combination of the other vectors in the set S .

- (b) Now we assume that at least one of the vectors in S is a linear combination of the others. For example, suppose that

$$\mathbf{v}_3 = \mathbf{v}_1 + 5\mathbf{v}_2. \quad (6.2)$$

Use vector algebra to rewrite Equation 6.2 so that $\mathbf{0}$ is expressed as a linear combination of \mathbf{v}_1 , \mathbf{v}_2 , and \mathbf{v}_3 such that the weight on \mathbf{v}_3 is not zero. Conclude that the set S is linearly dependent.

Now we provide a formal proof of Theorem 6.2, using the ideas from Activity 6.2.

Proof of Theorem 6.2. Let $S = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k\}$ be a set of vectors in \mathbb{R}^n . We will begin by verifying the first statement.

We assume that S is a linearly dependent set and show that at least one vector in S is a linear combination of the others. Since S is linearly dependent, there are scalars x_1, x_2, \dots, x_n , not all of which are 0, so that

$$x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + \cdots + x_k\mathbf{v}_k = \mathbf{0}. \quad (6.3)$$

We don't know which scalar(s) are not zero, but there is at least one. So let us assume that x_i is not zero for some i between 1 and k . We can then subtract $x_i\mathbf{v}_i$ from both sides of Equation (6.3) and divide by x_i to obtain

$$\mathbf{v}_i = \frac{x_1}{x_i}\mathbf{v}_1 + \frac{x_2}{x_i}\mathbf{v}_2 + \cdots + \frac{x_{i-1}}{x_i}\mathbf{v}_{i-1} + \frac{x_{i+1}}{x_i}\mathbf{v}_{i+1} + \frac{x_{i+2}}{x_i}\mathbf{v}_{i+2} + \cdots + \frac{x_k}{x_i}\mathbf{v}_k.$$

Thus, the vector \mathbf{v}_i is a linear combination of $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_{i-1}, \mathbf{v}_{i+1}, \dots, \mathbf{v}_k$, and at least one of the vectors in S is a linear combination of the other vectors in S .

To verify the second statement, we assume that at least one of the vectors in S can be written as a linear combination of the others and show that S is then a linearly dependent set. We don't know which vector(s) in S can be written as a linear combination of the others, but there is at least one. Let us suppose that \mathbf{v}_i is a linear combination of the vectors $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_{i-1}, \mathbf{v}_{i+1}, \dots, \mathbf{v}_k$ for some i between 1 and k . Then there exist scalars $x_1, x_2, \dots, x_{i-1}, x_{i+1}, \dots, x_n$ so that

$$\mathbf{v}_i = x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + \cdots + x_{i-1}\mathbf{v}_{i-1} + x_{i+1}\mathbf{v}_{i+1} + x_{i+2}\mathbf{v}_{i+2} + \cdots + x_k\mathbf{v}_k.$$

It follows that

$$\mathbf{0} = x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + \cdots + x_{i-1}\mathbf{v}_{i-1} + (-1)\mathbf{v}_i + x_{i+1}\mathbf{v}_{i+1} + x_{i+2}\mathbf{v}_{i+2} + \cdots + x_k\mathbf{v}_k.$$

So there are scalars x_1, x_2, \dots, x_n (with $x_i = -1$), not all of which are 0, so that

$$x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + \cdots + x_k\mathbf{v}_k = \mathbf{0}.$$

This makes S a linearly dependent set. ■

With a linearly dependent set, at least one of the vectors in the set is a linear combination of the others. With a linearly independent set, this cannot happen – no vector in the set can be written as a linear combination of the others. This result is given in the next theorem. You may be able to see how Theorems 6.2 and 6.3 are logically equivalent.



Theorem 6.3. A set $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k\}$ of vectors in \mathbb{R}^n is linearly independent if and only if no vector in the set can be written as a linear combination of the remaining vectors in the set.

Activity 6.3. As was hinted at in Preview Activity 6.1, an important consequence of a linearly independent set is that every vector in the span of the set can be written in one and only one way as a linear combination of vectors in the set. It is this uniqueness that makes linearly independent sets so useful. We explore this idea in this activity for a linearly independent set of three vectors. Let $S = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ be a linearly independent set of vectors in \mathbb{R}^n for some n , and let \mathbf{b} be a vector in $\text{Span } S$. To show that \mathbf{b} can be written in exactly one way as a linear combination of vectors in S , we assume that

$$\mathbf{b} = x_1 \mathbf{v}_1 + x_2 \mathbf{v}_2 + x_3 \mathbf{v}_3 \quad \text{and} \quad \mathbf{b} = y_1 \mathbf{v}_1 + y_2 \mathbf{v}_2 + y_3 \mathbf{v}_3$$

for some scalars x_1, x_2, x_3, y_1, y_2 , and y_3 . We need to demonstrate that $x_1 = y_1, x_2 = y_2$, and $x_3 = y_3$.

- (a) Use the two different ways of writing \mathbf{b} as a linear combination of $\mathbf{v}_1, \mathbf{v}_2$ and \mathbf{v}_3 to come up with a linear combination expressing $\mathbf{0}$ as a linear combination of these vectors.
- (b) Use the linear independence of the vectors $\mathbf{v}_1, \mathbf{v}_2$ and \mathbf{v}_3 to explain why $x_1 = y_1, x_2 = y_2$, and $x_3 = y_3$.

Activity 6.3 contains the general ideas to show that any vector in the span of a linearly independent set can be written in one and only one way as a linear combination of the vectors in the set. The weights of such a linear combination provide us a *coordinate system* for the vectors in terms of the basis. Two familiar examples of coordinate systems are the Cartesian coordinates in the xy -plane, and xyz -space. We will revisit the coordinate system idea in a later chapter.

In the next theorem we state and prove the general case of any number of linearly independent vectors producing unique representations as linear combinations.

Theorem 6.4. Let $S = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k\}$ be a linearly independent set of vectors in \mathbb{R}^n . Any vector in $\text{Span } S$ can be written in one and only one way as a linear combination of the vectors $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k$.

Proof. Let $S = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k\}$ be a linearly independent set of vectors in \mathbb{R}^n , and let \mathbf{b} be a vector in $\text{Span } S$. By definition, it follows that \mathbf{b} can be written as a linear combination of the vectors in S . It remains for us to show that this representation is unique. So assume that

$$\mathbf{b} = x_1 \mathbf{v}_1 + x_2 \mathbf{v}_2 + \cdots + x_k \mathbf{v}_k \quad \text{and} \quad \mathbf{b} = y_1 \mathbf{v}_1 + y_2 \mathbf{v}_2 + \cdots + y_k \mathbf{v}_k \quad (6.4)$$

for some scalars x_1, x_2, \dots, x_k , and y_1, y_2, \dots, y_k . Then

$$x_1 \mathbf{v}_1 + x_2 \mathbf{v}_2 + \cdots + x_k \mathbf{v}_k = y_1 \mathbf{v}_1 + y_2 \mathbf{v}_2 + \cdots + y_k \mathbf{v}_k.$$

Subtracting all terms from the right side and using a little vector algebra gives us

$$(x_1 - y_1) \mathbf{v}_1 + (x_2 - y_2) \mathbf{v}_2 + \cdots + (x_k - y_k) \mathbf{v}_k = \mathbf{0}.$$

The fact that S is a linearly independent set implies that

$$x_1 - y_1 = 0, x_2 - y_2 = 0, \dots, x_k - y_k = 0,$$

showing that $x_i = y_i$ for every i between 1 and k . We conclude that the representation of \mathbf{b} as a linear combination of the linearly independent vectors in S is unique. ■

Determining Linear Independence

The definition and our previous work give us a straightforward method for determining when a set of vectors in \mathbb{R}^n is linearly independent or dependent.

Activity 6.4. In this activity we learn how to use a matrix to determine in general if a set of vectors in \mathbb{R}^n is linearly independent or dependent. Suppose we have k vectors $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k$ in \mathbb{R}^n . To see if these vectors are linearly independent, we need to find the solutions to the vector equation

$$x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + \cdots + x_k\mathbf{v}_k = \mathbf{0}. \quad (6.5)$$

If we let $A = [\mathbf{v}_1 \ \mathbf{v}_2 \ \mathbf{v}_3 \ \cdots \ \mathbf{v}_k]$ and $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_k \end{bmatrix}$, then we can write the vector equation (6.5) in matrix form $A\mathbf{x} = \mathbf{0}$. Let B be the reduced row echelon form of A .

- (a) What can we say about the pivots of B in order for $A\mathbf{x} = \mathbf{0}$ to have exactly one solution? Under these conditions, are the vectors $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k$ linearly independent or dependent?
- (b) What can we say about the rows or columns of B in order for $A\mathbf{x} = \mathbf{0}$ to have infinitely many solutions? Under these conditions, are the vectors $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k$ linearly independent or dependent?

- (c) Use the result of parts (a) and (b) to determine if the vectors $\mathbf{v}_1 = \begin{bmatrix} 1 \\ -1 \\ 2 \\ 0 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} 1 \\ 0 \\ 2 \\ 3 \end{bmatrix}$,

and $\mathbf{v}_3 = \begin{bmatrix} 0 \\ 0 \\ 2 \\ 1 \end{bmatrix}$ in \mathbb{R}^4 are linearly independent or dependent. If dependent, write one of the vectors as a linear combination of the others. You may use the fact that the matrix

$$\begin{bmatrix} 1 & 1 & 0 \\ -1 & 0 & 0 \\ 2 & 2 & 2 \\ 0 & 3 & 1 \end{bmatrix} \text{ is row equivalent to } \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}.$$

Minimal Spanning Sets

It is important to note the differences and connections between linear independence, span, and minimal spanning set.

- The set $S = \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \right\}$ is not a minimal spanning set for \mathbb{R}^3 even though S is a

linearly independent set. Note that S does not span \mathbb{R}^3 since the vector $\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$ is not in $\text{Span } S$.

- The set $T = \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \right\}$ is not a minimal spanning set for \mathbb{R}^3 even though $\text{Span } T = \mathbb{R}^3$. Note that

$$\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix},$$

so T is not a linearly independent set.

- The set $U = \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\}$ is a minimal spanning set for \mathbb{R}^3 since it satisfies both characteristics of a minimal spanning set: $\text{Span } U = \mathbb{R}^3$ AND U is linearly independent.

The three concepts – linear independence, span, and minimal spanning set – are different. The important point to note is that minimal spanning set must be both linearly independent and span the space.

To find a minimal spanning set we will often need to find a smallest subset of a given set of vectors that has the same span as the original set of vectors. In this section we determine a method for doing so.

Activity 6.5. Let $\mathbf{v}_1 = \begin{bmatrix} -1 \\ 0 \\ 2 \end{bmatrix}$, $\mathbf{v}_2 = \begin{bmatrix} 2 \\ 0 \\ -4 \end{bmatrix}$, $\mathbf{v}_3 = \begin{bmatrix} 0 \\ 1 \\ 3 \end{bmatrix}$, and $\mathbf{v}_4 = \begin{bmatrix} -3 \\ 4 \\ 18 \end{bmatrix}$ in \mathbb{R}^3 . Assume that the reduced row echelon form of the matrix $A = \begin{bmatrix} -1 & 2 & 0 & -3 \\ 0 & 0 & 1 & 4 \\ 2 & -4 & 3 & 18 \end{bmatrix}$ is $\begin{bmatrix} 1 & -2 & 0 & 3 \\ 0 & 0 & 1 & 4 \\ 0 & 0 & 0 & 0 \end{bmatrix}$.

- (a) Write the general solution to the homogeneous system $A\mathbf{x} = \mathbf{0}$, where $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}$. Write

all linear combinations of \mathbf{v}_1 , \mathbf{v}_2 , \mathbf{v}_3 , and \mathbf{v}_4 that are equal to $\mathbf{0}$, using weights that only involve x_2 and x_4 .

- (b) Explain how we can conveniently choose the weights in the general solution to $A\mathbf{x} = \mathbf{0}$ to show that the vector \mathbf{v}_4 is a linear combination of \mathbf{v}_1 , \mathbf{v}_2 , and \mathbf{v}_3 . What does this tell us about $\text{Span}\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ and $\text{Span}\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4\}$?
- (c) Explain how we can conveniently choose the weights in the general solution to $A\mathbf{x} = \mathbf{0}$ to show why the vector \mathbf{v}_2 is a linear combination of \mathbf{v}_1 and \mathbf{v}_3 . What does this tell us about $\text{Span}\{\mathbf{v}_1, \mathbf{v}_3\}$ and $\text{Span}\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$?

(d) Is $\{\mathbf{v}_1, \mathbf{v}_3\}$ a minimal spanning set for $\text{Span}\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4\}$? Explain your response.

Activity 6.5 illustrates how we can use a matrix to determine a minimal spanning set for a given set of vectors $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k\}$ in \mathbb{R}^n .

- Form the matrix $A = [\mathbf{v}_1 \ \mathbf{v}_2 \ \cdots \ \mathbf{v}_k]$.
- Find the reduced row echelon form $[B \mid \mathbf{0}]$ of $[A \mid \mathbf{0}]$. If B contains non-pivot columns, say for example that the i th column is a non-pivot column, then we can choose the weight x_i corresponding to the i th column to be 1 and all weights corresponding to the other non-pivot columns to be 0 to make a linear combination of the columns of A that is equal to $\mathbf{0}$. This allows us to write \mathbf{v}_i as a linear combination of the vectors corresponding to the pivot columns of A as we did in the proof of Theorem 6.3. So every vector corresponding to a non-pivot column is in the span of the set of vectors corresponding to the pivot columns. The vectors corresponding to the pivot columns are linearly independent, since the matrix with those columns has every column as a pivot column. Thus, the set of vectors corresponding to the pivot columns of A forms a minimal spanning set for $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k\}$.

IMPORTANT NOTE! The set of pivot columns of the reduced row echelon form of A will normally not have the same span as the set of columns of A , so it is critical that we use columns of A , NOT B in our minimal spanning set.

Activity 6.6. Find a minimal spanning set for the span of the set

$$\left\{ \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 2 \\ 3 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 2 \\ 0 \end{bmatrix}, \begin{bmatrix} 4 \\ 1 \\ 0 \\ 0 \end{bmatrix} \right\}.$$

Activity 6.5 also illustrates a general process by which we can find a minimal spanning set – that is the smallest subset of vectors that has the same span. This process will be useful later when we consider vectors in arbitrary vector spaces. The idea is that if we can write one of the vectors in a set S as a linear combination of the remaining vectors, then we can remove that vector from the set and maintain the same span. In other words, begin with the span of a set S and follow these steps:

Step 1. If S is a linearly independent set, we already have a minimal spanning set.

Step 2. If S is not a linearly independent set, then one of the vectors in S is a linear combination of the others. Remove that vector from S to obtain a new set T . It will be the case that $\text{Span } T = \text{Span } S$.

Step 3. If T is a linearly independent set, then T is a minimal spanning set. If not, repeat steps 2 and 3 for the set T until you arrive at a linearly independent set.

This process is guaranteed to stop as long as the set contains at least one nonzero vector. A verification of the statement in Step 2 that $\text{Span } T = \text{Span } S$ is given in the next theorem.



Theorem 6.5. Let $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k\}$ be a set of vectors in \mathbb{R}^n so that for some i between 1 and k , \mathbf{v}_i is in $\text{Span}\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_{i-1}, \mathbf{v}_{i+1}, \dots, \mathbf{v}_k\}$. Then

$$\text{Span}\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k\} = \text{Span}\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_{i-1}, \mathbf{v}_{i+1}, \dots, \mathbf{v}_k\}.$$

Proof. Let $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k\}$ be a set of vectors in \mathbb{R}^n so that \mathbf{v}_i is in the span of $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_{i-1}, \mathbf{v}_{i+1}, \dots$, and \mathbf{v}_k for some i between 1 and k . To show that

$$\text{Span}\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k\} = \text{Span}\{\mathbf{v}_1, \dots, \mathbf{v}_{i-1}, \mathbf{v}_{i+1}, \dots, \mathbf{v}_k\},$$

we need to show that

- (1) every vector in $\text{Span}\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k\}$ is in $\text{Span}\{\mathbf{v}_1, \dots, \mathbf{v}_{i-1}, \mathbf{v}_{i+1}, \dots, \mathbf{v}_k\}$, and
- (2) every vector in $\text{Span}\{\mathbf{v}_1, \dots, \mathbf{v}_{i-1}, \mathbf{v}_{i+1}, \dots, \mathbf{v}_k\}$ is in $\text{Span}\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$.

Let us consider the second containment. Let \mathbf{x} be a vector in the span of $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_{i-1}, \mathbf{v}_{i+1}, \dots$, and \mathbf{v}_k . Then

$$\mathbf{x} = x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + \cdots + x_{i-1}\mathbf{v}_{i-1} + x_{i+1}\mathbf{v}_{i+1} + \cdots + x_k\mathbf{v}_k$$

for some scalars $x_1, x_2, \dots, x_{i-1}, x_{i+1}, \dots, x_k$. Note that

$$\mathbf{x} = x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + \cdots + x_{i-1}\mathbf{v}_{i-1} + (0)\mathbf{v}_i + x_{i+1}\mathbf{v}_{i+1} + \cdots + x_k\mathbf{v}_k$$

as well, so \mathbf{x} is in $\text{Span}\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k\}$. Thus,

$$\text{Span}\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_{i-1}, \mathbf{v}_{i+1}, \dots, \mathbf{v}_k\} \subseteq \text{Span}\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k\}.$$

(This same argument shows a more general statement that if S is a subset of T , then $\text{Span } S \subseteq \text{Span } T$.)

Now we demonstrate the first containment. Here we need the assumption that \mathbf{v}_i is in $\text{Span}\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_{i-1}, \mathbf{v}_{i+1}, \dots, \mathbf{v}_k\}$ for some i between 1 and k . That assumption gives us

$$\mathbf{v}_i = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_{i-1}\mathbf{v}_{i-1} + c_{i+1}\mathbf{v}_{i+1} + \cdots + c_k\mathbf{v}_k \quad (6.6)$$

for some scalars $c_1, c_2, \dots, c_{i-1}, c_{i+1}, \dots, c_k$. Now let \mathbf{x} be a vector in the span of $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k$. Then

$$\mathbf{x} = x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + \cdots + x_k\mathbf{v}_k$$

for some scalars x_1, x_2, \dots, x_k . Substituting from (6.6) shows that

$$\begin{aligned} \mathbf{x} &= x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + \cdots + x_k\mathbf{v}_k \\ &= x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + \cdots + x_{i-1}\mathbf{v}_{i-1} + x_i\mathbf{v}_i + x_{i+1}\mathbf{v}_{i+1} + \cdots + x_k\mathbf{v}_k \\ &= x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + \cdots + x_{i-1}\mathbf{v}_{i-1} \\ &\quad + x_i[c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_{i-1}\mathbf{v}_{i-1} + c_{i+1}\mathbf{v}_{i+1} + \cdots + c_k\mathbf{v}_k] \\ &\quad + x_{i+1}\mathbf{v}_{i+1} + \cdots + x_k\mathbf{v}_k \\ &= (x_1 + x_ic_1)\mathbf{v}_1 + (x_2 + x_ic_2)\mathbf{v}_2 + \cdots + (x_{i-1} + x_ic_{i-1})\mathbf{v}_{i-1} \\ &\quad + (x_{i+1} + x_ic_{i+1})\mathbf{v}_{i+1} + \cdots + (x_k + x_ic_k)\mathbf{v}_k. \end{aligned}$$



So \mathbf{x} is in $\text{Span}\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_{i-1}, \mathbf{v}_{i+1}, \dots, \mathbf{v}_k\}$ and

$$\text{Span}\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k\} \subseteq \text{Span}\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_{i-1}, \mathbf{v}_{i+1}, \dots, \mathbf{v}_k\}.$$

Since the two sets are subsets of each other, they must be equal sets. We conclude that

$$\text{Span}\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k\} = \text{Span}\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_{i-1}, \mathbf{v}_{i+1}, \dots, \mathbf{v}_k\}.$$

■

The result of Theorem 6.5 is that if we have a finite set S of vectors in \mathbb{R}^n , we can eliminate those vectors that are linear combinations of others until we obtain a smallest set of vectors that still has the same span. As mentioned earlier, we call such a minimal spanning set a basis.

Definition 6.6. Let S be a set of vectors in \mathbb{R}^n . A subset B of S is a **basis** for $\text{Span } S$ if B is linearly independent and $\text{Span } B = \text{Span } S$.

IMPORTANT NOTE: A basis is defined by two characteristics. A basis must span the space in question and a basis must be a linearly independent set. It is the linear independence that makes a basis a *minimal* spanning set.

We have worked with a familiar basis in \mathbb{R}^2 throughout our mathematical careers. A vector $\begin{bmatrix} a \\ b \end{bmatrix}$ in \mathbb{R}^2 can be written as

$$\begin{bmatrix} a \\ b \end{bmatrix} = a \begin{bmatrix} 1 \\ 0 \end{bmatrix} + b \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

So the set $\{\mathbf{e}_1, \mathbf{e}_2\}$, where $\mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $\mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ spans \mathbb{R}^2 . Since the columns of $[\mathbf{e}_1 \ \mathbf{e}_2]$ are linearly independent, so is the set $\{\mathbf{e}_1, \mathbf{e}_2\}$. Therefore, the set $\{\mathbf{e}_1, \mathbf{e}_2\}$ is a basis for \mathbb{R}^2 . The vector \mathbf{e}_1 is in the direction of the positive x -axis and the vector \mathbf{e}_2 is in the direction of the positive y -axis, so decomposing a vector $\begin{bmatrix} a \\ b \end{bmatrix}$ as a linear combination of \mathbf{e}_1 and \mathbf{e}_2 is akin to identifying the vector with the point (a, b) as we discussed earlier. The set $\{\mathbf{e}_1, \mathbf{e}_2\}$ is called the *standard basis* for \mathbb{R}^2 .

This idea is not restricted to \mathbb{R}^2 . Consider the vectors

$$\mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 0 \end{bmatrix}, \quad \mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \\ 0 \end{bmatrix}, \quad \dots, \quad \mathbf{e}_n = \begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}$$

in \mathbb{R}^n . That is, the vector \mathbf{e}_i is the vector with a 1 in the i th position and 0s everywhere else. Since the matrix $[\mathbf{e}_1 \ \mathbf{e}_2 \ \dots \ \mathbf{e}_n]$ is the identity matrix, the set $\{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$ is a basis for \mathbb{R}^n . The set $\{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$ is called the *standard basis* for \mathbb{R}^n .



As we will see later, bases¹ are of fundamental importance in linear algebra in that bases will allow us to define the dimension of a vector space and will provide us with coordinate systems.

We conclude this section with an important theorem that is similar Theorem 5.3.

Theorem 6.7. *Let A be an $m \times n$ matrix. The following statements are equivalent.*

- (1) *The matrix equation $A\mathbf{x} = \mathbf{b}$ has a unique solution for every vector \mathbf{b} in the span of the columns of A .*
- (2) *The matrix equation $A\mathbf{x} = \mathbf{0}$ has the unique solution $\mathbf{x} = \mathbf{0}$.*
- (3) *The columns of A are linearly independent.*
- (4) *The matrix A has a pivot position in each column.*

Examples

What follows are worked examples that use the concepts from this section.

Example 6.8. Let $\mathbf{v}_1 = \begin{bmatrix} 1 \\ 2 \\ 0 \\ 1 \end{bmatrix}$, $\mathbf{v}_2 = \begin{bmatrix} 0 \\ 6 \\ -1 \\ 5 \end{bmatrix}$, $\mathbf{v}_3 = \begin{bmatrix} 3 \\ -6 \\ 2 \\ -7 \end{bmatrix}$, and $\mathbf{v}_4 = \begin{bmatrix} 5 \\ -2 \\ 2 \\ -5 \end{bmatrix}$.

- (a) Is the set $S = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4\}$ linearly independent or dependent. If independent, explain why. If dependent, write one of the vectors in S as a linear combination of the other vectors in S .
- (b) Find a subset B of S that is a basis for $\text{Span } S$. Explain how you know you have a basis.

Example Solution.

- (a) We need to know the solutions to the vector equation

$$x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + x_3\mathbf{v}_3 + x_4\mathbf{v}_4 = \mathbf{0}.$$

If the equation has as its only solution $x_1 = x_2 = x_3 = x_4 = 0$ (the trivial solution), then the set S is linearly independent. Otherwise the set S is linearly dependent.

To find the solutions to this system, we row reduce the augmented matrix

$$\left[\begin{array}{cccc|c} 1 & 0 & 3 & 5 & 0 \\ 2 & 6 & -6 & -2 & 0 \\ 0 & -1 & 2 & 2 & 0 \\ 1 & 5 & -7 & -5 & 0 \end{array} \right].$$

¹The plural of basis is bases.

(Note that we really don't need the augmented column of zeros – row operations won't change that column at all. We just need to know that the column of zeros is there.) Technology shows that the reduced row echelon form of this augmented matrix is

$$\left[\begin{array}{cccc|c} 1 & 0 & 3 & 5 & 0 \\ 0 & 1 & -2 & -2 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array} \right].$$

The reduced row echelon form tells us that the vector equation is consistent, and the fact that there is no pivot in the fourth column shows that the system has a free variable and more than just the trivial solution. We conclude that S is linearly dependent.

Moreover, the general solution to our vector equation is

$$\begin{aligned} x_1 &= -3x_3 - 5x_4 \\ x_2 &= 2x_3 + 2x_4 \\ x_3 &\text{ is free} \\ x_4 &\text{ is free.} \end{aligned}$$

Letting $x_4 = 0$ and $x_3 = 1$ shows that one non-trivial solution to our vector equation is

$$x_1 = -3, \quad x_2 = 2, \quad x_3 = 1, \quad \text{and} \quad x_4 = 0.$$

Thus,

$$-3\mathbf{v}_1 + 2\mathbf{v}_2 + \mathbf{v}_3 = \mathbf{0},$$

or

$$\mathbf{v}_3 = 3\mathbf{v}_1 - 2\mathbf{v}_2$$

and we have written one vector in S as a linear combination of the other vectors in S .

- (b) We have seen that the pivot columns in a matrix A form a minimal spanning set (or basis) for the span of the columns of A . From part (a) we see that the pivot columns in the reduced row echelon form of $A = [\mathbf{v}_1 \mathbf{v}_2 \mathbf{v}_3 \mathbf{v}_4]$ are the first and second columns. So a basis for the span of the columns of A is $\{\mathbf{v}_1, \mathbf{v}_2\}$. Since the elements of S are the columns of A , we conclude that the set $B = \{\mathbf{v}_1, \mathbf{v}_2\}$ is a subset of S that is a basis for $\text{Span } S$.

Example 6.9. Let $\mathbf{v}_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$, $\mathbf{v}_2 = \begin{bmatrix} 3 \\ -7 \\ 2 \end{bmatrix}$, and $\mathbf{v}_3 = \begin{bmatrix} -5 \\ 6 \\ 10 \end{bmatrix}$.

- (a) Is the set $S = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ a basis for \mathbb{R}^3 ? Explain.

- (b) Let $\mathbf{v}_4 = \begin{bmatrix} -5 \\ 6 \\ h \end{bmatrix}$, where h is a scalar. Are there any values of h for which the set $S' = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_4\}$ is not a basis for \mathbb{R}^3 ? If so, find all such values of h and explain why S' is not a basis for \mathbb{R}^3 for those values of h .

Example Solution.



- (a) We need to know if the vectors in S are linearly independent and span \mathbb{R}^3 . Technology shows that the reduced row echelon form of

$$A = \begin{bmatrix} 1 & 3 & -5 \\ 1 & -7 & 6 \\ 0 & 2 & 10 \end{bmatrix}$$

is

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Since every column of $[\mathbf{v}_1 \ \mathbf{v}_2 \ \mathbf{v}_3]$ is a pivot column, the set $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is linearly independent. The fact that there is a pivot in every row of the matrix A means that the equation $A\mathbf{x} = \mathbf{b}$ is consistent for every \mathbf{b} in \mathbb{R}^3 . Since $A\mathbf{x}$ is a linear combination of the columns of A with weights from \mathbf{x} , it follows that the columns of A span \mathbb{R}^3 . We conclude that the set S is a basis for \mathbb{R}^3 .

- (b) Technology shows that a row echelon form of $A = [\mathbf{v}_1 \ \mathbf{v}_2 \ \mathbf{v}_4]$ is

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & -10 & 11 \\ 0 & 0 & h + \frac{11}{5} \end{bmatrix}.$$

The columns of A are all pivot columns (hence linearly independent) as long as $h \neq -\frac{11}{5}$, and are linearly dependent when $h = -\frac{11}{5}$. So the only value of h for which S' is not a basis for \mathbb{R}^3 is $h = -\frac{11}{5}$.

Summary

- A set $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k\}$ of vectors in \mathbb{R}^n is linearly independent if the vector equation

$$x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + \cdots + x_k\mathbf{v}_k = \mathbf{0}$$

for scalars x_1, x_2, \dots, x_k has only the trivial solution

$$x_1 = x_2 = x_3 = \cdots = x_k = 0.$$

Another way to think about this is that a set of vectors is linearly independent if no vector in the set can be written as a linear combination of the other vectors in the set.

- A set $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k\}$ of vectors in \mathbb{R}^n is linearly dependent if the vector equation

$$x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + \cdots + x_k\mathbf{v}_k = \mathbf{0}$$

has a nontrivial solution. That is, we can find scalars x_1, x_2, \dots, x_k that are not all 0 so that

$$x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + \cdots + x_k\mathbf{v}_k = \mathbf{0}.$$

Another way to think about this is that a set of vectors is linearly dependent if at least one vector in the set can be written as a linear combination of the other vectors in the set.



- If S is a set of vectors, a subset B of S is a basis for $\text{Span } S$ if B is a linearly independent set and $\text{Span } B = \text{Span } S$.
- Given a nonzero set S of vectors, we can remove vectors from S that are linear combinations of remaining vectors in S to obtain a linearly independent subset of S that has the same span as S .
- The columns of a matrix A are linearly independent if the equation $A\mathbf{x} = \mathbf{0}$ has only the trivial solution $\mathbf{x} = \mathbf{0}$.
- The set $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k\}$ is linearly independent if and only if every column of the matrix $A = [\mathbf{v}_1 \ \mathbf{v}_2 \ \mathbf{v}_3 \ \cdots \ \mathbf{v}_k]$, is a pivot column.
- If $A = [\mathbf{v}_1 \ \mathbf{v}_2 \ \mathbf{v}_3 \ \cdots \ \mathbf{v}_k]$, then the vectors in the pivot columns of A form a minimal spanning set for $\text{Span}\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k\}$.

Exercises

(1) Consider the following vectors in \mathbb{R}^3 :

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}, \quad \mathbf{v}_3 = \begin{bmatrix} 1 \\ 3 \\ 1 \end{bmatrix}$$

Is the set consisting of these vectors linearly independent? If so, explain why. If not, make a single change in one of the vectors so that the set is linearly independent.

(2) Consider the following vectors in \mathbb{R}^3 :

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix}, \quad \mathbf{v}_3 = \begin{bmatrix} 1 \\ 1 \\ c \end{bmatrix}$$

For which values of c is the set consisting of these vectors linearly independent?

(3) In a lab, there are three different water-benzene-acetic acid solutions: The first one with 36% water, 50% benzene and 14% acetic acid; the second one with 44% water, 46% benzene and 10% acetic acid; and the last one with 38% water, 49% benzene and 13% acid. Since the lab needs space, the lab coordinator wants to determine whether all solutions are needed, or if it is possible to create one of the solutions using the other two. Can you help the lab coordinator?

(4) Given vectors $\mathbf{v}_1 = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$ and $\mathbf{v}_2 = \begin{bmatrix} 0 \\ 2 \\ 1 \end{bmatrix}$, find a vector \mathbf{v}_3 in \mathbb{R}^3 so that the set consisting of \mathbf{v}_1 , \mathbf{v}_2 and \mathbf{v}_3 is linearly independent.



- (5) Consider the span of $S = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4\}$ where

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 4 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} 2 \\ 1 \\ 0 \\ 3 \end{bmatrix}, \mathbf{v}_3 = \begin{bmatrix} 3 \\ 2 \\ -1 \\ 1 \end{bmatrix}, \mathbf{v}_4 = \begin{bmatrix} 3 \\ 3 \\ 1 \\ 6 \end{bmatrix}.$$

- (a) Is the set S a minimal spanning set of $\text{Span } S$? If not, determine a minimal spanning set, i.e. a basis, of $\text{Span } S$.

- (b) Check that the vector $\mathbf{u} = \begin{bmatrix} 6 \\ 5 \\ -2 \\ 1 \end{bmatrix}$ is in $\text{Span } S$. Find the unique representation of \mathbf{u} in terms of the basis vectors.

- (6) Come up with a 4×3 matrix with linearly independent columns, if possible. If not, explain why not.

- (7) Come up with a 3×4 matrix with linearly independent columns, if possible. If not, explain why not.

- (8) Give an example of vectors $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ such that a minimal spanning set for $\text{Span}\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is equal to that of $\text{Span}\{\mathbf{v}_1, \mathbf{v}_2\}$; and an example of three vectors $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ such that a minimal spanning set for $\text{Span}\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is equal to that of $\text{Span}\{\mathbf{v}_1, \mathbf{v}_3\}$.

- (9) Label each of the following statements as True or False. Provide justification for your response.

- (a) **True/False** If $\mathbf{v}_1, \mathbf{v}_2$ and \mathbf{v}_3 are three vectors none of which is a multiple of another, then these vectors form a linearly independent set.

- (b) **True/False** If $\mathbf{v}_1, \mathbf{v}_2$ and \mathbf{v}_3 in \mathbb{R}^n are linearly independent vectors, then so are $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ and \mathbf{v}_4 for any \mathbf{v}_4 in \mathbb{R}^n .

- (c) **True/False** If $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ and \mathbf{v}_4 in \mathbb{R}^n are linearly independent vectors, then so are $\mathbf{v}_1, \mathbf{v}_2$ and \mathbf{v}_3 .

- (d) **True/False** A 3×4 matrix cannot have linearly independent columns.

- (e) **True/False** If two vectors span \mathbb{R}^2 , then they are linearly independent.

- (f) **True/False** The space \mathbb{R}^3 cannot contain four linearly independent vectors.

- (g) **True/False** If two vectors are linearly dependent, then one is a scalar multiple of the other.

- (h) **True/False** If a set of vectors in \mathbb{R}^n is linearly dependent, then the set contains more than n vectors.

- (i) **True/False** The columns of a matrix A are linearly independent if the equation $A\mathbf{x} = \mathbf{0}$ has only the trivial solution.



- (j) **True/False** Let $W = \text{Span}\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4\}$. If $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is a minimal spanning set for W , then $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_4\}$ cannot also be a minimal spanning set for W .
- (k) **True/False** Let $W = \text{Span}\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4\}$. If $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is a minimal spanning set for W , then $\{\mathbf{v}_1, \mathbf{v}_2\}$ cannot also be a minimal spanning set for W .
- (l) **True/False** If $\mathbf{v}_3 = 2\mathbf{v}_1 - 3\mathbf{v}_2$, then $\{\mathbf{v}_1, \mathbf{v}_2\}$ is a minimal spanning set for $\text{Span}\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$.

Project: Generating Bézier Curves

Bézier curves can be created as linear combinations of vectors. In this section we will investigate how cubic Bézier curves (the ones used for fonts) can be realized through linear and quadratic Bézier curves. We begin with linear Bézier curves.

Project Activity 6.1. Start with two vectors \mathbf{p}_0 and \mathbf{p}_1 . Linear Bézier curves are linear combinations

$$\mathbf{q} = (1 - t)\mathbf{p}_0 + t\mathbf{p}_1$$

of the vectors \mathbf{p}_0 and \mathbf{p}_1 for scalars t between 0 and 1. (You can visualize these linear combinations using the GeoGebra file Linear Bezier at <https://www.geogebra.org/m/HvrPhh86>. With this file you can draw the vectors \mathbf{q} for varying values of t . You can move the points \mathbf{p}_0 and \mathbf{p}_1 in the GeoGebra file, and the slider controls the values of t . The point identified with \mathbf{q} is traced as t is changed.) For this activity, we will see what the curve \mathbf{q} corresponds to by evaluating certain points on the curve in a specific example. Let $\mathbf{p}_0 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$ and $\mathbf{p}_1 = \begin{bmatrix} 6 \\ 3 \end{bmatrix}$.

- (a) What are the components of the vector $(1 - t)\mathbf{p}_0 + t\mathbf{p}_1$ if $t = \frac{1}{2}$? Where is this vector in relation to \mathbf{p}_0 and \mathbf{p}_1 ? Explain.
- (b) What are the components of the vector $(1 - t)\mathbf{p}_0 + t\mathbf{p}_1$ if $t = \frac{1}{3}$? Where is this vector in relation to \mathbf{p}_0 and \mathbf{p}_1 ? Explain.
- (c) What are the components of the vector $(1 - t)\mathbf{p}_0 + t\mathbf{p}_1$ for an arbitrary t ? Where is this vector in relation to \mathbf{p}_0 and \mathbf{p}_1 ? Explain.

For each value of t , the vector $\mathbf{q} = (1 - t)\mathbf{p}_0 + t\mathbf{p}_1$ is a linear combination of the vectors \mathbf{p}_0 and \mathbf{p}_1 . Note that when $t = 0$, we have $\mathbf{q} = \mathbf{p}_0$ and when $t = 1$ we have $\mathbf{q} = \mathbf{p}_1$, and for $0 \leq t \leq 1$ Project Activity 6.1 shows that the vectors \mathbf{q} trace out the line segment from \mathbf{p}_0 to \mathbf{p}_1 . The span $\{(1 - t)\mathbf{p}_0 + t\mathbf{p}_1\}$ of the vectors \mathbf{p}_0 and \mathbf{p}_1 for $0 \leq t \leq 1$ is a linear Bézier curve. Once we have a construction like this, it is natural in mathematics to extend it and see what happens. We do that in the next activity to construct quadratic Bézier curves.

Project Activity 6.2. Let \mathbf{p}_0 , \mathbf{p}_1 , and \mathbf{p}_2 be vectors in the plane. We can then let

$$\mathbf{q}_0 = (1 - t)\mathbf{p}_0 + t\mathbf{p}_1 \quad \text{and} \quad \mathbf{q}_1 = (1 - t)\mathbf{p}_1 + t\mathbf{p}_2$$

be the linear Bézier curves as defined in Project Activity 6.1. Since \mathbf{q}_0 and \mathbf{q}_1 are vectors, we can define \mathbf{r} as

$$\mathbf{r} = (1 - t)\mathbf{q}_0 + t\mathbf{q}_1.$$



(You can visualize these linear combinations using the GeoGebra file Quadratic Bezier at <https://www.geogebra.org/m/VWCZZBXZ>. With this file you can draw the vectors \mathbf{r} for varying values of t . You can move the points \mathbf{p}_0 , \mathbf{p}_1 , and \mathbf{p}_2 in the GeoGebra file, and the slider controls the values of t . The point identified with \mathbf{r} is traced as t is changed.) In this activity we investigate how the vectors \mathbf{r} change as t changes. For the remainder of this activity, let $\mathbf{p}_0 = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$, $\mathbf{p}_1 = \begin{bmatrix} 8 \\ 4 \end{bmatrix}$, and $\mathbf{p}_2 = \begin{bmatrix} 6 \\ -3 \end{bmatrix}$.

- At what point (in terms of \mathbf{p}_0 , \mathbf{p}_1 , and \mathbf{p}_2) is the vector $\mathbf{r} = (1-t)\mathbf{q}_0 + t\mathbf{q}_1$ when $t = 0$? Explain using the definition of \mathbf{r} .
- At what point (in terms of \mathbf{p}_0 , \mathbf{p}_1 , and \mathbf{p}_2) is the vector $\mathbf{r} = (1-t)\mathbf{q}_0 + t\mathbf{q}_1$ when $t = 1$? Explain using the definition of \mathbf{r} .
- Find by hand the components of the vector $(1-t)\mathbf{q}_0 + t\mathbf{q}_1$ with $t = \frac{1}{4}$. Compare with the result of the GeoGebra file.

The span $\{(1-t)\mathbf{q}_0 + t\mathbf{q}_1\}$ of the vectors \mathbf{q}_0 and \mathbf{q}_1 , or the set of points traced out by the vectors \mathbf{r} for $0 \leq t \leq 1$, is a quadratic Bézier curve. To understand why this curve is called quadratic, we examine the situation in a general context in the following activity.

Project Activity 6.3. Let \mathbf{p}_0 , \mathbf{p}_1 , and \mathbf{p}_2 be arbitrary vectors in the plane. Write $\mathbf{r} = (1-t)\mathbf{q}_0 + t\mathbf{q}_1$ as a linear combination of \mathbf{p}_0 , \mathbf{p}_1 , and \mathbf{p}_2 . That is, write \mathbf{r} in the form $a_0\mathbf{p}_0 + a_1\mathbf{p}_1 + a_2\mathbf{p}_2$ for some scalars (that may depend on t) a_0 , a_1 , and a_2 . Explain why the result leads us to call these vectors *quadratic* Bézier curves.

Notice that if any one of the \mathbf{p}_i lies on the line determined by the other two vectors, then the quadratic Bézier curve is just a line segment. So to obtain something non-linear we need to choose our vectors so that that doesn't happen.

Quadratic Bézier curves are limited, because their graphs are parabolas. For applications we need higher order Bézier curves. In the next activity we consider cubic Bézier curves.

Project Activity 6.4. Start with four vectors \mathbf{p}_0 , \mathbf{p}_1 , \mathbf{p}_2 , \mathbf{p}_3 – the points defined by these vectors are called *control points* for the curve. As with the linear and quadratic Bézier curves, we let

$$\mathbf{q}_0 = (1-t)\mathbf{p}_0 + t\mathbf{p}_1, \quad \mathbf{q}_1 = (1-t)\mathbf{p}_1 + t\mathbf{p}_2, \quad \text{and} \quad \mathbf{q}_2 = (1-t)\mathbf{p}_2 + t\mathbf{p}_3.$$

Then let

$$\mathbf{r}_0 = (1-t)\mathbf{q}_0 + t\mathbf{q}_1 \quad \text{and} \quad \mathbf{r}_1 = (1-t)\mathbf{q}_1 + t\mathbf{q}_2.$$

We take this one step further to generate the cubic Bézier curves by letting

$$\mathbf{s} = (1-t)\mathbf{r}_0 + t\mathbf{r}_1.$$

(You can visualize these linear combinations using the GeoGebra file Cubic Bezier at <https://www.geogebra.org/m/EDAhudy9>. With this file you can draw the vectors \mathbf{s} for varying values of t . You can move the points \mathbf{p}_0 , \mathbf{p}_1 , \mathbf{p}_2 , and \mathbf{p}_3 in the GeoGebra file, and the slider controls the values of t . The point identified with \mathbf{s} is traced as t is changed.) In this activity we



investigate how the vectors \mathbf{s} change as t changes. For the remainder of this activity, let $\mathbf{p}_0 = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$, $\mathbf{p}_1 = \begin{bmatrix} 4 \\ 5 \end{bmatrix}$, $\mathbf{p}_2 = \begin{bmatrix} 9 \\ -3 \end{bmatrix}$, and $\mathbf{p}_3 = \begin{bmatrix} 2 \\ 0 \end{bmatrix}$.

- At what point (in terms of \mathbf{p}_0 , \mathbf{p}_1 , \mathbf{p}_2 , and \mathbf{p}_3) is the vector $\mathbf{s} = (1 - t)\mathbf{r}_0 + t\mathbf{r}_1$ when $t = 0$? Explain using the definition of \mathbf{s} .
- At what point (in terms of \mathbf{p}_0 , \mathbf{p}_1 , \mathbf{p}_2 , and \mathbf{p}_3) is the vector $\mathbf{s} = (1 - t)\mathbf{r}_0 + t\mathbf{r}_1$ when $t = 1$? Explain using the definition of \mathbf{s} .
- Find by hand the components of the vector $(1 - t)\mathbf{r}_0 + t\mathbf{r}_1$ with $t = \frac{3}{4}$. Compare with the result of the GeoGebra file.

The span $\{(1 - t)\mathbf{r}_0 + t\mathbf{r}_1\}$ of the vectors \mathbf{r}_0 and \mathbf{r}_1 , or the set of points traced out by the vectors \mathbf{s} for $0 \leq t \leq 1$, is a cubic Bézier curve. To understand why this curve is called cubic, we examine the situation in a general context in the following activity.

Project Activity 6.5. Let \mathbf{p}_0 , \mathbf{p}_1 , \mathbf{p}_2 , and \mathbf{p}_3 be arbitrary vectors in the plane. Write $\mathbf{s} = (1 - t)\mathbf{r}_0 + t\mathbf{r}_1$ as a linear combination of \mathbf{p}_0 , \mathbf{p}_1 , \mathbf{p}_2 , and \mathbf{p}_3 . That is, write \mathbf{s} in the form $b_0\mathbf{p}_0 + b_1\mathbf{p}_1 + b_2\mathbf{p}_2 + b_3\mathbf{p}_3$ for some scalars (that may depend on t) b_0 , b_1 , b_2 , and b_3 . Explain why the result leads us to call these vectors *cubic* Bézier curves.

Just as with the quadratic case, we need certain subsets of the set of control vectors to be linearly independent so that the cubic Bézier curve does not degenerate to a quadratic or linear Bézier curve.

More complicated and realistic shapes can be represented by piecing together two or more Bézier curves as illustrated with the letter “S” in Figure 6.1. Suppose we have two cubic Bézier curves, the first with control points \mathbf{p}_0 , \mathbf{p}_1 , \mathbf{p}_2 , and \mathbf{p}_3 and the second with control points \mathbf{p}'_0 , \mathbf{p}'_1 , \mathbf{p}'_2 , and \mathbf{p}'_3 . You may have noticed that \mathbf{p}_1 lies on the tangent line to the first Bézier curve at \mathbf{p}_0 and that \mathbf{p}_2 lies on the tangent line to the first Bézier curve at \mathbf{p}_3 . (Play around with the program Cubic Bezier to convince yourself of these statements. This can be proved in a straightforward manner using vector calculus.) So if we want to make a smooth curve from these two Bézier curves, the curves will need to join together smoothly at \mathbf{p}_3 and \mathbf{p}'_0 . This will force $\mathbf{p}_3 = \mathbf{p}'_0$ and the tangents at $\mathbf{p}_3 = \mathbf{p}'_0$ will have to match. This implies that \mathbf{p}_2 , \mathbf{p}_3 , and \mathbf{p}'_1 all have to lie on this common tangent line. Keeping this idea in mind, use the GeoGebra file Cubic Bezier Pair at <https://www.geogebra.org/m/UwxQ6RPk> to find control points for the pair of Bézier curves that create your own letter S.

Section 7

Matrix Transformations

Focus Questions

By the end of this section, you should be able to give precise and thorough answers to the questions listed below. You may want to keep these questions in mind to focus your thoughts as you complete the section.

- What is a matrix transformation?
- What properties do matrix transformations have? (In particular, what properties make matrix transformations *linear*?)
- What is the domain of a matrix transformation defined by an $m \times n$ matrix? Why?
- What are the range and codomain of a matrix transformation defined by an $m \times n$ matrix? Why?
- What does it mean for a matrix transformation to be one-to-one? If T is a matrix transformation represented as $T(\mathbf{x}) = A\mathbf{x}$, what are the conditions on A that make T a one-to-one transformation?
- What does it mean for a matrix transformation to be onto? If T is a matrix transformation represented as $T(\mathbf{x}) = A\mathbf{x}$, what are the conditions on A that make T an onto transformation?

Application: Computer Graphics

As we will discuss, left multiplication by an $m \times n$ matrix defines a function from \mathbb{R}^n to \mathbb{R}^m . Such a function defined by matrix multiplication is called a matrix transformation. In this section we study some of the properties of matrix transformations and understand how, using the pivots of the matrix, to determine when the output of a matrix transformation covers the whole space \mathbb{R}^m or when a transformation maps distinct vectors to distinct outputs.

Matrix transformations are used extensively in computer graphics to produce animations as seen

in video games and movies. For example, consider the dancing figure at left in Figure 7.1. We can identify certain control points (e.g., the point at the neck, where the arms join the torso, etc.) to mark the locations of important points. Using just the control points we can reconstruct the figure. Each control point can be represented as a vector, and so we can manipulate the figure by manipulating the control points with matrix transformations. We will explore this idea in more detail later in this section.

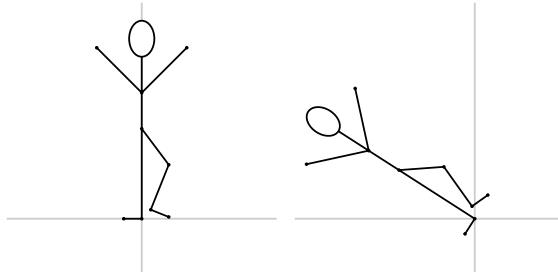


Figure 7.1: A dancing figure and a rotated dancing figure.

Introduction

In this section we will consider special functions which take vectors as inputs and produce vectors as outputs. We will use matrix multiplication to produce the output vectors.

If A is an $m \times n$ matrix and \mathbf{x} is a vector in \mathbb{R}^n , then the matrix-vector product $A\mathbf{x}$ is a vector in \mathbb{R}^m . (Pick some specific n, m values to understand this statement better.) Therefore, left multiplication by the matrix A takes an input vector \mathbf{x} in \mathbb{R}^n and produces an output vector $A\mathbf{x}$ in \mathbb{R}^m , which we will refer to as the *image* of \mathbf{x} under the transformation. This defines a function T from \mathbb{R}^n to \mathbb{R}^m where

$$T(\mathbf{x}) = A\mathbf{x}.$$

These functions are the matrix transformations.

Definition 7.1. A **matrix transformation** is a function $T : \mathbb{R}^n \rightarrow \mathbb{R}^m$ defined by

$$T(\mathbf{x}) = A\mathbf{x}$$

for some $m \times n$ matrix A .

Many of the transformations we consider in this section are from \mathbb{R}^2 to \mathbb{R}^2 so that we can visualize the transformations. As an example, let us consider the transformation T defined by

$$T \left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \right) = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}.$$

If we plot the input vectors $\mathbf{u}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, $\mathbf{u}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$, $\mathbf{u}_3 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$, and $\mathbf{u}_4 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$ (as (blue) circles) and their images $T(\mathbf{u}_1) = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, $T(\mathbf{u}_2) = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ -1 \end{bmatrix}$,

$T(\mathbf{u}_3) = \begin{bmatrix} 1 \\ -2 \end{bmatrix}$, and $T(\mathbf{u}_4) = \begin{bmatrix} -1 \\ -1 \end{bmatrix}$ (as (red) \times 's) on the same set of axes as shown in Figure 7.2, we see that this transformation reflects the input vectors across the x -axis. We can also see this algebraically since the reflection of the point (x_1, x_2) around the x -axis is the point $(x_1, -x_2)$, and

$$T\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \begin{bmatrix} x_1 \\ -x_2 \end{bmatrix}.$$

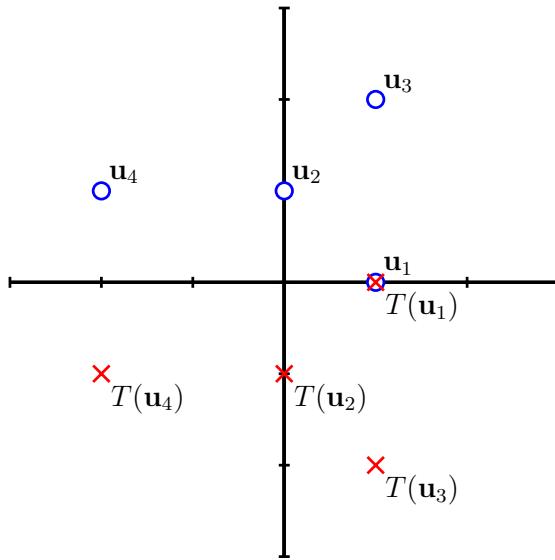


Figure 7.2: Inputs and outputs of the transformation T .

Preview Activity 7.1. We now consider other transformations from \mathbb{R}^2 to \mathbb{R}^2 .

- (1) Suppose a transformation T is defined by

$$T\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}.$$

- (a) Find $T(\mathbf{u}_i)$ for each of $\mathbf{u}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, $\mathbf{u}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$, $\mathbf{u}_3 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$, and $\mathbf{u}_4 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$. (In other words, substitute $\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3, \mathbf{u}_4$ into the formula above to see what output is obtained.)
- (b) Plot all input vectors and their images on the same axes in \mathbb{R}^2 . Clearly identify which image corresponds to which input vector. Then give a geometric description of what this transformation does.

- (2) The transformation in the introduction performs a reflection across the x -axis. Find a matrix transformation that performs a reflection across the y -axis.

(3) Suppose a transformation T is defined by

$$T(\mathbf{x}) = A\mathbf{x},$$

where

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}.$$

- (a) Find $T(\mathbf{u}_i)$ for each of $\mathbf{u}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, $\mathbf{u}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$, $\mathbf{u}_3 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$, and $\mathbf{u}_4 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$.
- (b) Plot all input vectors and their images on the same axes in \mathbb{R}^2 . Give a geometric description of this transformation.
- (c) Is there an input vector which produces $\mathbf{b} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ as an output vector?
- (d) Find all input vectors that produce the output vector $\mathbf{b} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$. Is there a unique input vector, or multiple input vectors?

Properties of Matrix Transformations

A matrix transformation is a function. When dealing with functions in previous mathematics courses we have used the terms domain and range with our functions. Recall that the domain of a function is the set of all allowable inputs into the function and the range of a function is the set of all outputs of the function. We do the same with transformations. If T is the matrix transformation $T(\mathbf{x}) = A\mathbf{x}$ for some $m \times n$ matrix A , then T maps vectors from \mathbb{R}^n into \mathbb{R}^m . So \mathbb{R}^n is the *domain* of T – the set of all input vectors. However, the set \mathbb{R}^m is only the target set for T and not necessarily the range of T . We call \mathbb{R}^m the *codomain* of T , while the *range* of T is the set of all output vectors. The range is always a subset of the codomain, but the two sets do not have to be equal. In addition, if a vector \mathbf{b} in \mathbb{R}^m satisfies $\mathbf{b} = T(\mathbf{x})$ for some \mathbf{x} in \mathbb{R}^n , then we say that \mathbf{b} is the *image* of \mathbf{x} under the transformation T .

Because of the properties of the matrix-vector product, if the matrix transformation T is defined by $T(\mathbf{x}) = A\mathbf{x}$ for some $m \times n$ matrix A , then

$$T(\mathbf{u} + \mathbf{v}) = A(\mathbf{u} + \mathbf{v}) = A\mathbf{u} + A\mathbf{v}$$

and

$$T(c\mathbf{u}) = A(c\mathbf{u}) = cA\mathbf{u} = cT(\mathbf{u})$$

for any vectors \mathbf{u} and \mathbf{v} in \mathbb{R}^n and any scalar c . So every matrix transformation T satisfies the following two important properties:

- (1) $T(\mathbf{u} + \mathbf{v}) = T(\mathbf{u}) + T(\mathbf{v})$ and
- (2) $T(c\mathbf{u}) = cT(\mathbf{u})$.

The first property says that a matrix transformation T preserves sums of vectors and the second that T preserves scalar multiples of vectors.



Activity 7.1. Let T be a matrix transformation, and let \mathbf{u} and \mathbf{v} be vectors in the domain of T so that $T(\mathbf{u}) = \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}$ and $T(\mathbf{v}) = \begin{bmatrix} -3 \\ 1 \\ 4 \end{bmatrix}$.

- (a) Exactly which vector is $T(2\mathbf{u} - 3\mathbf{v})$? Explain.
- (b) If a and b are any scalars, what is the vector $T(a\mathbf{u} + b\mathbf{v})$? Why?

As we saw in Activity 7.1, we can combine the two properties of a matrix transformation T into one: for any scalars a and b and any vectors \mathbf{u} and \mathbf{v} in the domain of T we have

$$T(a\mathbf{u} + b\mathbf{v}) = aT(\mathbf{u}) + bT(\mathbf{v}). \quad (7.1)$$

We can then extend equation (7.1) (by mathematical induction) to any finite linear combination of vectors. That is, if $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k$ are any vectors in the domain of a matrix transformation T and if x_1, x_2, \dots, x_k are any scalars, then

$$T(x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + \cdots + x_k\mathbf{v}_k) = x_1T(\mathbf{v}_1) + x_2T(\mathbf{v}_2) + \cdots + x_kT(\mathbf{v}_k). \quad (7.2)$$

In other words, a matrix transformation preserves linear combinations. For this reason matrix transformations are examples of a larger set of transformation that are called *linear* transformations. We will discuss general linear transformations in a later section.

There is one other important property of a matrix transformation for us to consider. The functions we encountered in earlier mathematics courses, e.g., $f(x) = 2x + 1$, could send the input 0 to any output. However, as a consequence of the definition, any matrix transformation T maps the zero vector to the zero vector because

$$T(\mathbf{0}) = A\mathbf{0} = \mathbf{0}.$$

Note that the two vectors $\mathbf{0}$ in the last equation may not be the same vector – if $T : \mathbb{R}^n \rightarrow \mathbb{R}^m$, then the first $\mathbf{0}$ is in \mathbb{R}^n and the second in \mathbb{R}^m . It should be clear from the context which vector $\mathbf{0}$ is meant.

Onto and One-to-One Transformations

The problems we have been asking about solutions to systems of linear equations can be rephrased in terms of matrix transformations. The question about whether a system $A\mathbf{x} = \mathbf{b}$ is consistent for any vector \mathbf{b} is also a question about the existence of a vector \mathbf{x} so that $T(\mathbf{x}) = \mathbf{b}$, where T is the matrix transformation defined by $T(\mathbf{x}) = A\mathbf{x}$.

Activity 7.2. Let T be the matrix transformation defined by $T(\mathbf{x}) = A\mathbf{x}$ where A is

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 2 \end{bmatrix}.$$

- (a) Find $T\left(\begin{bmatrix} 1 \\ 1 \end{bmatrix}\right)$ and $T\left(\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}\right)$. If it is not possible to find one or both of the output vectors, indicate why.
- (b) What are the domain and codomain of T ? Why? (Recall that the domain is the space of all input vectors, while the codomain is the space in which the output vectors are contained.)
- (c) Can you find a vector \mathbf{x} for which $T(\mathbf{x}) = \begin{bmatrix} 2 \\ 3 \\ 6 \end{bmatrix}$? Can you find a vector \mathbf{x} for which $T(\mathbf{x}) = \begin{bmatrix} 2 \\ 3 \\ 1 \end{bmatrix}$?
- (d) Which $\mathbf{b} = \begin{bmatrix} a \\ b \\ c \end{bmatrix}$ are the image vectors for this transformation? Is the range of T equal to the codomain of T ? Explain.
- (e) The previous question can be rephrased as a matrix equation question. We are asking whether $A\mathbf{x} = \mathbf{b}$ is consistent for every \mathbf{b} . How is the answer to this question related to the pivots of A ?

If T is a matrix transformation, Activity 7.2 illustrates that the range of a matrix transformation T may not equal its codomain. In other words, there may be vectors \mathbf{b} in the codomain of T that are not the image of any vector in the domain of T . If it is the case for a matrix transformation T that there is always a vector \mathbf{x} in the domain of T such that $T(\mathbf{x}) = \mathbf{b}$ for any vector \mathbf{b} in the codomain of T , then T is given a special name.

Definition 7.2. A matrix transformation T from \mathbb{R}^n to \mathbb{R}^m is **onto** if each \mathbf{b} in \mathbb{R}^m is the image of *at least one* \mathbf{x} in \mathbb{R}^n .

So the matrix transformation T from \mathbb{R}^n to \mathbb{R}^m defined by $T(\mathbf{x}) = A\mathbf{x}$ is onto if the equation $A\mathbf{x} = \mathbf{b}$ has a solution for each vector \mathbf{b} in \mathbb{R}^m . Since the vectors $A\mathbf{x}$ are linear combinations of the columns of A , T is onto exactly when the span of the columns of A is all of \mathbb{R}^m . Activity 7.2 shows us that T is onto if every row of A contains a pivot.

Another question to ask about matrix transformations is how many vectors there can be that map onto a given output vector.

Activity 7.3. Let T be the matrix transformation defined by $T(\mathbf{x}) = A\mathbf{x}$ where A is

$$\begin{bmatrix} 1 & 3 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

- (a) Find $T\left(\begin{bmatrix} 1 \\ 1 \end{bmatrix}\right)$ and $T\left(\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}\right)$. If it is not possible to find one or both of the output vectors, indicate why.



- (b) What are the domain and codomain of T ? Why?
- (c) Find $T \left(\begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix} \right)$. Are there any other \mathbf{x} 's for which $T(\mathbf{x})$ is this same output vector?
 (Hint: Set up an equation to solve for such \mathbf{x} 's.)
- (d) Assume more generally that for some vector \mathbf{b} , there is a vector \mathbf{x} so that $T(\mathbf{x}) = \mathbf{b}$. Write this as a matrix equation to determine how many solutions this equation has. Explain. How is the answer to this question related to the pivots of A ?

The uniqueness of a solution to $A\mathbf{x} = \mathbf{b}$ is the same as saying that the matrix transformation T defined by $T(\mathbf{x}) = A\mathbf{x}$ maps exactly one vector to \mathbf{b} . A matrix transformation T that has the property that every image vector is an image in exactly one way is also a special type of transformation.

Definition 7.3. A matrix transformation T from \mathbb{R}^n to \mathbb{R}^m is **one-to-one** if each \mathbf{b} in \mathbb{R}^m is the image of *at most* one \mathbf{x} in \mathbb{R}^n .

So the matrix transformation T from \mathbb{R}^n to \mathbb{R}^m defined by $T(\mathbf{x}) = A\mathbf{x}$ is one-to-one if the equation $A\mathbf{x} = \mathbf{b}$ has a unique solution whenever $A\mathbf{x} = \mathbf{b}$ is consistent. Since the vectors $A\mathbf{x}$ are linear combinations of the columns of A , the unique solution requirement indicates that any output vector can be written in exactly one way as a linear combination of the columns of A . This implies that the columns of A are linearly independent. Activity 7.3 indicates that this happens when every column of A is a pivot column.

To summarize, if T is a matrix transformation defined by $T(\mathbf{x}) = A\mathbf{x}$, then T is onto if every row of A contains a pivot, and T is one-to-one if every column of A is a pivot column. It is important to note the difference: being one-to-one depends on the rows of A and being onto depends on the columns of A .

Having a matrix transformation from \mathbb{R}^n to \mathbb{R}^m can tell us things about m and n . For example, when a matrix transformation from \mathbb{R}^n to \mathbb{R}^m is one-to-one, it means that there is a unique input vector for every output vector. Since a matrix transformation preserves the algebraic structure of \mathbb{R}^n , this implies that the collection of the images of the vectors in the domain of T form a copy of \mathbb{R}^n inside of \mathbb{R}^m . If we think of T as a one-to-one matrix transformation with $T(\mathbf{x}) = A\mathbf{x}$ for some $m \times n$ matrix, then every column of A will have to be a pivot column. It follows that if there is a one-to-one matrix transformation from \mathbb{R}^n to \mathbb{R}^m , we must have $m \geq n$. Similarly, if a matrix transformation T from \mathbb{R}^n to \mathbb{R}^m is onto, then for each \mathbf{b} in \mathbb{R}^m , if we select one vector in the domain of T whose image is \mathbf{b} , then the collection of these vectors in the domain of T is a copy of \mathbb{R}^m inside of \mathbb{R}^n . So if there is an onto matrix transformation from \mathbb{R}^n to \mathbb{R}^m , then $n \geq m$. As a consequence, the only way a matrix transformation from \mathbb{R}^n to \mathbb{R}^m is both one-to-one and onto is if $n = m$.

We conclude this section by adding new equivalent conditions to Theorems 5.3 and 6.7 from Sections 5 and 6.

Theorem 7.4. Let A be an $m \times n$ matrix. The following statements are equivalent.

- (1) The matrix equation $A\mathbf{x} = \mathbf{b}$ has a solution for every vector \mathbf{b} in \mathbb{R}^m .
- (2) Every vector \mathbf{b} in \mathbb{R}^m can be written as a linear combination of the columns of A .



- (3) The span of the columns of A is \mathbb{R}^m .
- (4) The matrix A has a pivot position in each row.
- (5) The matrix transformation T from \mathbb{R}^n to \mathbb{R}^m defined by $T(\mathbf{x}) = A\mathbf{x}$ is onto.

Theorem 7.5. Let A be an $m \times n$ matrix. The following statements are equivalent.

- (1) The matrix equation $A\mathbf{x} = \mathbf{b}$ has a unique solution for every vector \mathbf{b} in the span of the columns of A .
- (2) The matrix equation $A\mathbf{x} = \mathbf{0}$ has the unique solution $\mathbf{x} = \mathbf{0}$.
- (3) The columns of A are linearly independent.
- (4) The matrix A has a pivot position in each column.
- (5) The matrix transformation T from \mathbb{R}^n to \mathbb{R}^m defined by $T(\mathbf{x}) = A\mathbf{x}$ is one-to-one.

We will continue to add to these theorems, which will eventually give us many different but equivalent perspectives to look at a linear algebra problem. Please keep these equivalent criteria in mind when considering the best possible approach to a problem.

Examples

What follows are worked examples that use the concepts from this section.

Example 7.6. Let $A = \begin{bmatrix} 1 & 1 & -1 & -1 \\ 3 & 6 & 0 & 3 \\ 2 & -1 & -5 & -8 \end{bmatrix}$ and let $T(\mathbf{x}) = A\mathbf{x}$.

- (a) Identify the domain of T . Explain your reasoning.
- (b) Is T one-to-one. Explain.
- (c) Is T onto? If yes, explain why. If no, describe the range of T as best you can, both algebraically and graphically.

Example Solution.

- (a) Since A is a 3×4 matrix, A has four columns. Now $A\mathbf{x}$ is a linear combination of the columns of A with weights from \mathbf{x} , so \mathbf{x} must have four entries to correspond to the columns of A . We conclude that the domain of T is \mathbb{R}^4 .
- (b) Technology shows that the reduced row echelon form of A is

$$\left[\begin{array}{cccc} 1 & 0 & -2 & -3 \\ 0 & 1 & 1 & 2 \\ 0 & 0 & 0 & 0 \end{array} \right].$$

Since A contains non-pivot columns, the homogeneous system $A\mathbf{x} = \mathbf{0}$ has infinitely many solutions. So T is not one-to-one. In other words, if there is a column of A that is a non-pivot column, then A is not one-to-one.



- (c) Since the reduced row echelon form of A has rows of zeros, there will be vectors \mathbf{b} in \mathbb{R}^3 such that the reduced row echelon form of $[A \mathbf{b}]$ will have a row of the form $[0 \ 0 \ 0 \ c]$ for some nonzero scalar c . This means that $T(\mathbf{x}) = A\mathbf{x} = \mathbf{b}$ will have no solution and T is not onto. In other words, if there is a row of A that does not contain a pivot, then T is not onto.

- (d) To determine the vectors $\mathbf{b} = \begin{bmatrix} r \\ s \\ t \end{bmatrix}$ so that $T(\mathbf{x}) = A\mathbf{x} = \mathbf{b}$ is consistent, we row reduce the augmented matrix $[A \mid \mathbf{b}]$. Technology shows that an echelon form of $[A \mathbf{b}]$ is

$$\left[\begin{array}{cccc|c} 1 & 1 & -1 & -1 & r \\ 0 & 3 & 3 & 6 & s - 3r \\ 0 & 0 & 0 & 0 & t - 5r + s \end{array} \right].$$

Thus, the system $A\mathbf{x} = \mathbf{b}$ is consistent if and only if $-5r + s + t = 0$. We can then write the general output vector to this system as

$$\mathbf{b} = \begin{bmatrix} r \\ s \\ 5r - s \end{bmatrix} = r \begin{bmatrix} 1 \\ 0 \\ 5 \end{bmatrix} + s \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix},$$

with r and s any scalars. Since there are two free variables, the vectors \mathbf{b} in \mathbb{R}^3 define a plane through the origin. Letting $r = 0$ and $s = 1$ and $r = 1$ and $s = 0$, we see that two points that lie on this plane are $(0, 1, -1)$ and $(1, 0, 5)$. So the range of T is the plane through the origin and the points $(0, 1, -1)$ and $(1, 0, 5)$.

Example 7.7. A matrix transformation $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ defined by

$$T \left(\begin{bmatrix} x \\ y \end{bmatrix} \right) = \begin{bmatrix} cx \\ y \end{bmatrix}$$

is a contraction in the x direction if $0 < c < 1$ and a dilation in the x direction if $c > 1$.

- (a) Find a matrix A such that $T(\mathbf{x}) = A\mathbf{x}$.

- (b) Sketch the square S with vertices $\mathbf{u}_1 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$, $\mathbf{u}_2 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, $\mathbf{u}_3 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, and $\mathbf{u}_4 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$. Determine and sketch the image of S under T if $c = 2$.

Example Solution.

- (a) Since

$$\begin{bmatrix} c & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} cx \\ y \end{bmatrix},$$

the matrix $A = \begin{bmatrix} c & 0 \\ 0 & 1 \end{bmatrix}$ has the property that $T(\mathbf{x}) = A\mathbf{x}$.



- (b) We can determine the image of S under T by calculating what T does to the vertices of S . Notice that

$$T(\mathbf{u}_1) = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$T(\mathbf{u}_2) = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 2 \\ 0 \end{bmatrix}$$

$$T(\mathbf{u}_3) = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

$$T(\mathbf{u}_4) = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

Since T is a linear map, the image of S under T is the polygon with vertices $(0, 0)$, $(1, 0)$, $(2, 1)$, and $(0, 1)$ as shown in Figure 7.3. From Figure 7.3 we can see that T stretches the figure in the x direction only by a factor of 2.

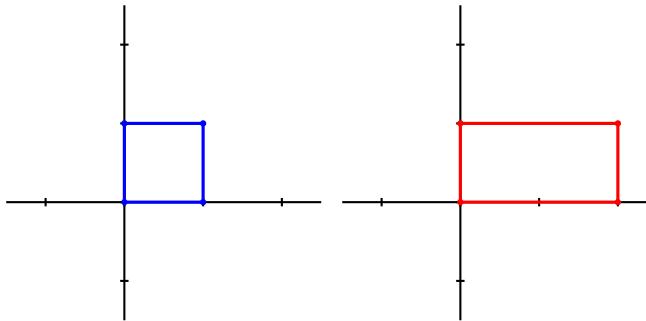


Figure 7.3: The input square S and the output $T(S)$.

Summary

In this section we determined how to represent any matrix transformation from \mathbb{R}^n to \mathbb{R}^m as a matrix transformation, and what it means for a matrix transformation to be one-to-one and onto.

- A matrix transformation is a function $T : \mathbb{R}^n \rightarrow \mathbb{R}^m$ defined by $T(\mathbf{x}) = A\mathbf{x}$ for some $m \times n$ matrix A .
- A matrix transformation T from \mathbb{R}^n to \mathbb{R}^m satisfies

$$T(a\mathbf{u} + b\mathbf{v}) = aT(\mathbf{u}) + bT(\mathbf{v})$$

for any scalars a and b and any vectors \mathbf{u} and \mathbf{v} in \mathbb{R}^n . The fact that T preserves linear combinations is why we say that T is a linear transformation.

- An $m \times n$ matrix A defines the matrix transformation T via

$$T(\mathbf{x}) = A\mathbf{x}.$$

The domain of this transformation is \mathbb{R}^n because the matrix-vector product $A\mathbf{x}$ is only defined if \mathbf{x} is an $n \times 1$ vector.

- If A is an $m \times n$ matrix, then the codomain of the matrix transformation T defined by $T(\mathbf{x}) = A\mathbf{x}$ is \mathbb{R}^m . This is because the matrix-vector product $A\mathbf{x}$ with \mathbf{x} an $n \times 1$ vector is an $m \times 1$ vector. The range of T is the subset of the codomain of T consisting of all vectors of the form $T(\mathbf{x})$ for vectors \mathbf{x} in the domain of T .
- A matrix transformation T from \mathbb{R}^n to \mathbb{R}^m is **one-to-one** if each \mathbf{b} in \mathbb{R}^m is the image of *at most* one \mathbf{x} in \mathbb{R}^n . If T is a matrix transformation represented as $T(\mathbf{x}) = A\mathbf{x}$, then T is one-to-one if each column of A is a pivot column, or if the columns of A are linearly independent.
- A matrix transformation T from \mathbb{R}^n to \mathbb{R}^m is **onto** if each \mathbf{b} in \mathbb{R}^m is the image of *at least one* \mathbf{x} in \mathbb{R}^n . If T is a matrix transformation represented as $T(\mathbf{x}) = A\mathbf{x}$, then T is onto if each row of A contains a pivot position, or if the span of the columns of A is all of \mathbb{R}^m .

Exercises

- (1) Given matrix $A = \begin{bmatrix} 1 & 2 & 1 \\ 1 & 0 & -3 \end{bmatrix}$, write the coordinate form of the transformation T defined by $T(\mathbf{x}) = A\mathbf{x}$. (Note: Writing a transformation in coordinate form refers to writing the transformation in terms of the entries of the input and output vectors.)
- (2) Suppose the transformation T is defined by $T(\mathbf{x}) = A\mathbf{x}$ where

$$A = \begin{bmatrix} 1 & 1 & -1 \\ 2 & 1 & 1 \\ 4 & 1 & 4 \end{bmatrix}.$$

Determine if $\mathbf{b} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$ is in the range of T . If so, find all \mathbf{x} 's which map to \mathbf{b} .

- (3) Suppose T is a matrix transformation and

$$T(\mathbf{v}_1) = \begin{bmatrix} 1 \\ 2 \end{bmatrix}, T(\mathbf{v}_2) = \begin{bmatrix} -2 \\ 3 \end{bmatrix}$$

Find $T(2\mathbf{v}_1 - 5\mathbf{v}_2)$.

- (4) Given a matrix transformation defined as

$$T \left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \right) = \begin{bmatrix} 2x_1 - x_3 \\ -x_1 + 2x_2 + x_3 \\ 3x_2 - 4x_3 \end{bmatrix}$$

determine the matrix A for which $T(\mathbf{x}) = A\mathbf{x}$.



- (5) Suppose a matrix transformation T defined by $T(\mathbf{x}) = A\mathbf{x}$ for some unknown A matrix satisfies

$$T\left(\begin{bmatrix} 1 \\ 0 \end{bmatrix}\right) = \begin{bmatrix} 2 \\ 1 \end{bmatrix} \text{ and } T\left(\begin{bmatrix} 0 \\ 1 \end{bmatrix}\right) = \begin{bmatrix} -1 \\ 3 \end{bmatrix}.$$

Use the matrix transformation properties to determine $T(\mathbf{x})$ where $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$. Use the expression for $T(\mathbf{x})$ to determine the matrix A .

- (6) For each of the following matrices, determine if the transformation T defined by $T(\mathbf{x}) = A\mathbf{x}$ is onto and if T is one-to-one.

$$(a) A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & -3 \end{bmatrix}$$

$$(b) A = \begin{bmatrix} 1 & 1 & 2 \\ 2 & 2 & 4 \end{bmatrix}$$

$$(c) A = \begin{bmatrix} 1 & 1 & 2 \\ 1 & 2 & 3 \\ -1 & 1 & 2 \end{bmatrix}$$

$$(d) A = \begin{bmatrix} 1 & 1 \\ 2 & 3 \\ 3 & 0 \end{bmatrix}$$

- (7) Come up with an example of a one-to-one transformation from \mathbb{R}^3 to \mathbb{R}^4 , if possible. If not, explain why not.
- (8) Come up with an example of an onto transformation from \mathbb{R}^3 to \mathbb{R}^4 , if possible. If not, explain why not.
- (9) Come up with an example of a one-to-one but not onto transformation from \mathbb{R}^4 to \mathbb{R}^4 , if possible. If not, explain why not.
- (10) Two students are talking about when a matrix transformation is one-to-one.

Student 1: If we have a matrix transformation, then we need to check that $A\mathbf{x} = \mathbf{b}$ has a unique solution for every \mathbf{b} for which $A\mathbf{x} = \mathbf{b}$ has a solution, right?

Student 2: Well, that's the definition. Each \mathbf{b} in the codomain has to be the image of at most one \mathbf{x} in the domain. So when \mathbf{b} is in the range, corresponding to $A\mathbf{x} = \mathbf{b}$ having a solution, then there is exactly one solution \mathbf{x} .

Student 1: But wouldn't it be enough to check that $A\mathbf{x} = \mathbf{0}$ has a unique solution? Doesn't that translate to the other \mathbf{b} vectors? If there is a unique solution for one \mathbf{b}_1 , then there can't be infinitely many solutions for another \mathbf{b}_2 .

Student 2: I don't know. It feels to me as if changing the right hand side could change whether there is a unique solution, or infinitely many solutions, or no solution.



Which part of the above conversation do you agree with? Which parts need fixing?

- (11) Show that if T is a matrix transformation from \mathbb{R}^n to \mathbb{R}^m and L is a line in \mathbb{R}^n , then $T(L)$, the image of L , is a line or a single vector. (Note that a line in \mathbb{R}^n is the set of all vectors of the form $\mathbf{v} + c\mathbf{w}$ where c is a scalar, and \mathbf{v}, \mathbf{w} are two fixed vectors in \mathbb{R}^n .)
- (12) Label each of the following statements as True or False. Provide justification for your response.
- (a) **True/False** The range of a transformation is the same as the codomain of the transformation.
 - (b) **True/False** The codomain of a transformation T defined by $T(\mathbf{x}) = A\mathbf{x}$ is the span of the columns of A .
 - (c) **True/False** A one-to-one transformation is a transformation where each input has a unique output.
 - (d) **True/False** A one-to-one transformation is a transformation where each output can only come from a unique input.
 - (e) **True/False** If a matrix transformation from \mathbb{R}^n to \mathbb{R}^n is one-to-one, then it is also onto.
 - (f) **True/False** A matrix transformation from \mathbb{R}^2 to \mathbb{R}^3 cannot be onto.
 - (g) **True/False** A matrix transformation from \mathbb{R}^3 to \mathbb{R}^2 cannot be onto.
 - (h) **True/False** A matrix transformation from \mathbb{R}^3 to \mathbb{R}^2 cannot be one-to-one.
 - (i) **True/False** If the columns of a matrix A are linearly independent, then the transformation T defined by $T(\mathbf{x}) = A\mathbf{x}$ is onto.
 - (j) **True/False** If the columns of a matrix A are linearly independent, then the transformation T defined by $T(\mathbf{x}) = A\mathbf{x}$ is one-to-one.
 - (k) **True/False** If A is an $m \times n$ matrix with n pivots, then the transformation $\mathbf{x} \mapsto A\mathbf{x}$ is onto.
 - (l) **True/False** If A is an $m \times n$ matrix with n pivots, then the transformation $\mathbf{x} \mapsto A\mathbf{x}$ is one-to-one.
 - (m) **True/False** If \mathbf{u} is in the range of a matrix transformation T , then there is an \mathbf{x} in the domain of T such that $T(\mathbf{x}) = \mathbf{u}$.
 - (n) **True/False** If T is a one-to-one matrix transformation, then $T(\mathbf{x}) = \mathbf{0}$ has a non-trivial solution.
 - (o) **True/False** If the transformations $T_1 : \mathbb{R}^m \rightarrow \mathbb{R}^n$ and $T_2 : \mathbb{R}^n \rightarrow \mathbb{R}^p$ are onto, then the transformation $T_2 \circ T_1$ defined by $T_2 \circ T_1(\mathbf{x}) = T_2(T_1(\mathbf{x}))$ is also onto.
 - (p) **True/False** If the transformations $T_1 : \mathbb{R}^m \rightarrow \mathbb{R}^n$ and $T_2 : \mathbb{R}^n \rightarrow \mathbb{R}^p$ are one-to-one, then the transformation $T_2 \circ T_1$ defined by $T_2 \circ T_1(\mathbf{x}) = T_2(T_1(\mathbf{x}))$ is also one-to-one.



Project: The Geometry of Matrix Transformations

In this section we will consider certain types of matrix transformations and analyze their geometry. Much more would be needed for real computer graphics, but the essential ideas are contained in our examples. A GeoGebra applet is available at <https://www.geogebra.org/m/rh4bzxee> for you to use to visualize the transformations in this project.

Project Activity 7.1. We begin with transformations that produce the rotated dancing image in Figure 7.1. Let R be the matrix transformation from \mathbb{R}^2 to \mathbb{R}^2 defined by

$$R\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}.$$

These matrices are the rotation matrices.

- (a) Suppose $\theta = \frac{\pi}{2}$. Then

$$R\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}.$$

- i. Find the images of $\mathbf{u}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, $\mathbf{u}_2 = \begin{bmatrix} \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} \end{bmatrix}$, and $\mathbf{u}_3 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ under R .
- ii. Plot the points determined by the vectors from part i. The matrix transformation R performs a rotation. Based on this small amount of data, what would you say the angle of rotation is for this transformation R ?

- (b) Now let R be the general matrix transformation defined by the matrix

$$\begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix}.$$

Follow the steps indicated to show that R performs a counterclockwise rotation of an angle θ around the origin. Let P be the point defined by the vector $\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \cos(\alpha) \\ \sin(\alpha) \end{bmatrix}$ and Q the point defined by the vector $\begin{bmatrix} w \\ z \end{bmatrix} = \begin{bmatrix} \cos(\alpha + \theta) \\ \sin(\alpha + \theta) \end{bmatrix}$ as illustrated in Figure 7.4.

- i. Use the angle sum trigonometric identities

$$\begin{aligned} \cos(A + B) &= \cos(A)\cos(B) - \sin(A)\sin(B) \\ \sin(A + B) &= \cos(A)\sin(B) + \cos(B)\sin(A) \end{aligned}$$

to show that

$$\begin{aligned} w &= \cos(\theta)x - \sin(\theta)y \\ z &= \sin(\theta)x + \cos(\theta)y. \end{aligned}$$

- ii. Now explain why the counterclockwise rotation around the origin by an angle θ can be represented by left multiplication by the matrix

$$\begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix}.$$



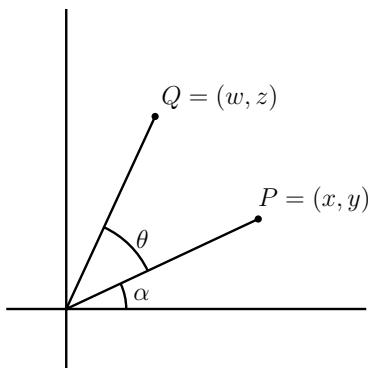


Figure 7.4: A rotation in the plane.

Project Activity 7.1 presented the rotation matrices. Other matrices have different effects.

Project Activity 7.2. Different matrix transformations

- (a) Let S be the matrix transformation from \mathbb{R}^2 to \mathbb{R}^2 defined by

$$S \left(\begin{bmatrix} x \\ y \end{bmatrix} \right) = \begin{bmatrix} 1 & 0.5 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}.$$

Determine the entries of the output vector $S \left(\begin{bmatrix} x \\ y \end{bmatrix} \right)$ and explain the action of the transformation S on the dancing figure as illustrated in Figure 7.5. (The transformation S is called a *shear* in the x direction.)

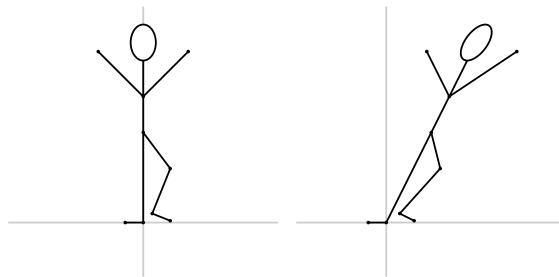


Figure 7.5: A dancing figure and a sheared dancing figure.

- (b) Let C be the matrix transformation from \mathbb{R}^2 to \mathbb{R}^2 defined by

$$C \left(\begin{bmatrix} x \\ y \end{bmatrix} \right) = \begin{bmatrix} 0.65 & 0 \\ 0 & 0.65 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}.$$

Determine the entries of the output vector $C \left(\begin{bmatrix} x \\ y \end{bmatrix} \right)$ and explain the action of the transformation C on the dancing figure as illustrated in Figure 7.6. (The transformation C is

called a *contraction*.) How would your response change if each 0.65 was changed to 2 in the matrix C ?

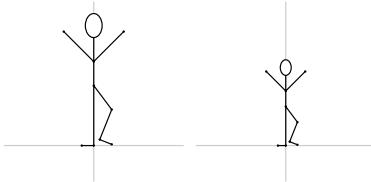


Figure 7.6: A dancing figure and a contracted dancing figure.

So far we have seen specific matrix transformations perform rotations, shears, and contractions. We can combine these, and other, matrix transformations by composition to change figures in different ways, and to create animations of geometric figures. (As we will see later, combining transformations needs to be done carefully in order to obtain the result we want. For example, if we want to first rotate then translate, in what order should the matrices be applied?)



Part II

Matrices

Section 8

Matrix Operations

Focus Questions

By the end of this section, you should be able to give precise and thorough answers to the questions listed below. You may want to keep these questions in mind to focus your thoughts as you complete the section.

- Under what conditions can we add two matrices and how is the matrix sum defined?
- Under what conditions can we multiply a matrix by a scalar and how is a scalar multiple of a matrix defined?
- Under what conditions can we multiply two matrices and how is the matrix product defined?
- What properties do matrix addition, scalar multiplication of matrices and matrix multiplication satisfy? Are these properties similar to properties that are satisfied by vector operations?
- What are two properties that make matrix multiplication fundamentally different than our standard product of real numbers?
- What is the interpretation of matrix multiplication from the perspective of linear transformations?
- How is the transpose of a matrix defined?

Application: Algorithms for Matrix Multiplication

Matrix multiplication is widely used in applications ranging from scientific computing and pattern recognition to counting paths in graphs. As a consequence, much work is being done in developing efficient algorithms for matrix multiplication.

We will see that a matrix product can be calculated through the row-column method. Recall

that the product of two 2×2 matrices $A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$ and $B = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix}$ is given by

$$AB = \begin{bmatrix} a_{11}b_{11} + a_{12}b_{21} & a_{11}b_{12} + a_{12}b_{22} \\ a_{21}b_{11} + a_{22}b_{21} & a_{21}b_{12} + a_{22}b_{22} \end{bmatrix},$$

This product involves eight scalar multiplications and some scalar additions. As we will see, multiplication is more computationally expensive than addition, so we will focus on multiplication. In 1969, a German mathematician named Volker Strassen showed¹ that the product of two 2×2 matrices can be calculated using only seven multiplications. While this is not much of an improvement, the Strassen algorithm can be applied to larger matrices, using matrix partitions (which allow for parallel computation), and its publication led to additional research on faster algorithms for matrix multiplication. More details are provided later in this section.

Introduction

A vector is a list of numbers in a specified order and a matrix is an ordered array of objects. In fact, a vector can be thought of as a matrix of size $n \times 1$. Vectors and matrices are so alike in this way that it would seem natural that we can define operations on matrices just as we did with vectors.

Recall that a matrix is made of rows and columns – the entries reading from left to right form the *rows* of the matrix and the entries reading from top to bottom form the *columns*. The number of rows and columns of a matrix is called the *size* of the matrix, so an $m \times n$ matrix has m rows and n columns. If we label the entry in the i th row and j th column of a matrix A as a_{ij} , then we write $A = [a_{ij}]$.

We can generalize the operations of addition and scalar multiplication on vectors to matrices similarly. Given two matrices $A = [a_{ij}]$ and $B = [b_{ij}]$ of the same size, we define the sum $A + B$ by

$$A + B = [a_{ij} + b_{ij}]$$

when the sizes of the matrices A and B match. In other words, for matrices of the same size the matrix addition is defined by adding corresponding entries in the matrices. For example,

$$\begin{bmatrix} 1 & 2 \\ -2 & 3 \end{bmatrix} + \begin{bmatrix} 1 & 3 \\ 2 & 4 \end{bmatrix} = \begin{bmatrix} 2 & 5 \\ 0 & 7 \end{bmatrix}.$$

We define the scalar multiple of a matrix $A = [a_{ij}]$ by scalar c to be the matrix cA defined by

$$cA = [ca_{ij}].$$

This means that we multiply each entry of the matrix A by the scalar c . As an example,

$$3 \begin{bmatrix} 1 & 2 \\ -2 & 3 \end{bmatrix} = \begin{bmatrix} 3 & 6 \\ -6 & 9 \end{bmatrix}.$$

Even though we did not have a multiplication operation on vectors, we had a matrix-vector product, which is a special case of a matrix-matrix product since a vector is a matrix with one column. However, generalizing the matrix-vector product to a matrix-matrix product is not immediate

¹Strassen, Volker, Gaussian Elimination is not Optimal, Number. Math. 13, p. 354-356, 1969



as it is not immediately clear what we can do with the other columns. We will consider this question in this section.

Note that all of the matrix operations can be performed on a calculator. After entering each matrix in the calculator, just use $+$, $-$ and \times operations to find the result of the matrix operation. (Just for fun, try using \div with matrices to see if it will work.)

Preview Activity 8.1.

- (1) Pick three different varying sizes of pairs of A, B matrices which can be added. For each pair:
 - (a) Find the matrices $A + B$ and $B + A$.
 - (b) How are the two matrices $A + B$ and $B + A$ related? What does this tell us about matrix addition?
- (2) Let $A = \begin{bmatrix} 1 & 0 \\ -2 & 8 \end{bmatrix}$, $B = \begin{bmatrix} 1 & 1 \\ 3 & 4 \end{bmatrix}$, and $C = \begin{bmatrix} 0 & -5 \\ 1 & 6 \end{bmatrix}$. Determine the entries of the matrix $A + 2B - 7C$.
- (3) Now we turn to multiplication of matrices. Our first goal is to find out what conditions we need on the sizes of matrices A and B if the matrix-matrix product AB is defined and what the size of the resulting product matrix is. We know the condition and the size of the result in the special case of B being a vector, i.e., a matrix with one column. So our conjectures for the general case should match what we know in the special case.

In each part of this problem, use any appropriate tool (e.g., your calculator, Maple, Mathematica, Wolfram|Alpha) to determine the matrix product AB , if it exists. If you obtain a product, write it down and explain how its size is related to the sizes of A and B . If you receive an error, write down the error and guess why the error occurred and/or what it means.

$$(a) A = \begin{bmatrix} 1 & 2 & 0 \\ 0 & 1 & 1 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 3 & 5 & 0 \\ 0 & -2 & 1 \end{bmatrix}$$

$$(b) A = \begin{bmatrix} 1 & 2 & 0 \\ 0 & 1 & 1 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 3 & 0 \\ 5 & -2 \\ 0 & 1 \end{bmatrix}$$

$$(c) A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 0 & 1 \\ 0 & 2 & 0 \end{bmatrix}$$

$$(d) A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \\ 7 & 8 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 1 & 2 & 3 \\ -1 & 1 & 1 \end{bmatrix}$$

- (e) Make a guess for the condition on the sizes of two matrices A, B for which the product AB is defined. How is the size of the product matrix related to the sizes of A and B ?

- (4) The final matrix products, when defined, in problem 3 might seem unrelated to the individual matrices at first. In this problem, we will uncover this relationship using our knowledge of the matrix-vector product.

Let $A = \begin{bmatrix} 3 & -1 \\ -2 & 3 \end{bmatrix}$ and $B = \begin{bmatrix} 0 & 2 & 1 \\ 1 & 3 & 2 \end{bmatrix}$.

- Calculate AB using any tool.
- Using the matrix-vector product, calculate Ax where x is the first column (i.e., calculate $A \begin{bmatrix} 0 \\ 1 \end{bmatrix}$), and then the second column of B (i.e., calculate $A \begin{bmatrix} 2 \\ 3 \end{bmatrix}$), and then the third column of B (i.e., calculate $A \begin{bmatrix} 1 \\ 2 \end{bmatrix}$). Do you notice these output vectors within AB ?
- Describe as best you can a definition of AB using the matrix-vector product.

Properties of Matrix Addition and Multiplication by Scalars

Just as we were able to define an algebra of vectors with addition and multiplication by scalars, we can define an algebra of matrices. We will see that the properties of these operations on matrices are immediate generalizations of the properties of the operations on vectors. We will then see how the matrix product arises through the connection of matrices to linear transformations. Finally, we define the transpose of a matrix. The transpose of a matrix will be useful in applications such as graph theory and least-squares fitting of curves, as well as in advanced topics as inner product spaces and the dual space of a vector space.

We learned in Preview Activity 8.1 that we can add two matrices of the same size together by adding corresponding entries and we can multiply any matrix by a scalar by multiplying each entry of the matrix by that scalar. More generally, if $A = [a_{ij}]$ and $B = [b_{ij}]$ are $m \times n$ matrices and c is any scalar, then

$$A + B = [a_{ij} + b_{ij}] \quad \text{and} \quad cA = [ca_{ij}].$$

As we have done each time we have introduced a new operation, we ask what properties the operation has. For example, you determined in Preview Activity 8.1 that addition of matrices is a commutative operation. More specifically, for every two $m \times n$ matrices A and B , $A + B = B + A$. We can use similar arguments to verify the following properties of matrix addition and multiplication by scalars. Notice that these properties are very similar to the properties of addition and scalar multiplication of vectors we discussed earlier. This should come as no surprise since the n -dimensional vectors are $n \times 1$ matrices. In a strange twist, we will see that matrices themselves can be considered as vectors when we discuss vector spaces in a later section.

Theorem 8.1. *Let A , B , and C be $m \times n$ matrices and let a and b be scalars. Then*

- (1) $A + B = B + A$ (this property tells us that matrix addition is commutative)
- (2) $(A + B) + C = A + (B + C)$ (this property tells us that matrix addition is associative)



- (3) The $m \times n$ matrix 0 whose entries are all 0 has the property that $A + 0 = A$. The matrix 0 is called the **zero matrix** (It is generally clear from the context what the size of the 0 matrix is.).
- (4) The scalar multiple $(-1)A$ of the matrix A has the property that $(-1)A + A = 0$. The matrix $(-1)A = -A$ is called the **additive inverse** of the matrix A .
- (5) $(a + b)A = aA + bA$ (this property tells us that scalar multiplication of matrices distributes over scalar addition)
- (6) $a(A + B) = aA + aB$ (this property tells us that scalar multiplication of matrices distributes over matrix addition)
- (7) $(ab)A = a(bA)$
- (8) $1A = A$.

Later on, we will see that these properties define the set of all $m \times n$ matrices as a *vector space*. These properties just say that, regarding addition and multiplication by scalars, we can manipulate matrices just as we do real numbers. Note, however, we have not yet defined an operation of multiplication on matrices. That is the topic for the next section.

A Matrix Product

As we saw in Preview Activity 8.1, a matrix-matrix product can be found in a way which makes use of and also generalizes the matrix-vector product.

Definition 8.2. The **matrix product** of a $k \times m$ matrix A and an $m \times n$ matrix $B = [\mathbf{b}_1 \ \mathbf{b}_2 \ \cdots \ \mathbf{b}_n]$ with columns $\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_n$ is the $k \times n$ matrix

$$[A\mathbf{b}_1 \ A\mathbf{b}_2 \ \cdots \ A\mathbf{b}_n].$$

We now consider the motivation behind this definition by thinking about the matrix transformations corresponding to each of the matrices A, B and AB . Recall that left multiplication by an $m \times n$ matrix B defines a transformation T from \mathbb{R}^n to \mathbb{R}^m by $T(\mathbf{x}) = B\mathbf{x}$. The domain of T is \mathbb{R}^n because the number of components of \mathbf{x} have to match the number of entries in each of row of B in order for the matrix-vector product $B\mathbf{x}$ to be defined. Similarly, a $k \times m$ matrix A defines a transformation A from \mathbb{R}^m to \mathbb{R}^k . Since transformations are functions, we can compose them as long as the output vectors of the inside transformation lie in the domain of the outside transformation. Therefore if T is the inside transformation and S is the outside transformation, the composition $S \circ T$ is defined. So a natural question to ask is if we are given

- a transformation T from \mathbb{R}^n to \mathbb{R}^m where $T(\mathbf{x}) = B\mathbf{x}$ for an $m \times n$ matrix B and
- a transformation S from \mathbb{R}^m to \mathbb{R}^k with $S(\mathbf{y}) = A\mathbf{y}$ for some $k \times m$ matrix A ,

is there a matrix that represents the transformation $S \circ T$ defined by $(S \circ T)(\mathbf{x}) = S(T(\mathbf{x}))$? We investigate this question in the next activity in the special case of a 2×3 matrix A and a 3×2 matrix B .



Activity 8.1. In this activity, we look for the meaning of the matrix product from a transformation perspective. Let S and T be matrix transformations defined by

$$S(\mathbf{y}) = A\mathbf{y} \quad \text{and} \quad T(\mathbf{x}) = B\mathbf{x},$$

where

$$A = \begin{bmatrix} 1 & 2 & 0 \\ 0 & 1 & 1 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 3 & 0 \\ 5 & -2 \\ 0 & 1 \end{bmatrix}.$$

- (a) What are the domains and codomains of S and T ? Why is the composite transformation $S \circ T$ defined? What is the domain of $S \circ T$? What is the codomain of $S \circ T$? (Recall that $S \circ T$ is defined by $(S \circ T)(\mathbf{x}) = S(T(\mathbf{x}))$, i.e., we substitute the output $T(\mathbf{x})$ as the input into the transformation S .)
- (b) Let $\mathbf{x} = \begin{bmatrix} x \\ y \end{bmatrix}$. Determine the components of $T(\mathbf{x})$.
- (c) Find the components of $S \circ T(\mathbf{x}) = S(T(\mathbf{x}))$.
- (d) Find a matrix C so that $S(T(\mathbf{x})) = C\mathbf{x}$.
- (e) Use the definition of composition of transformations and the definitions of the S and T transformations to explain why it is reasonable to define AB to be the matrix C . Does the matrix C agree with the

$$AB = \begin{bmatrix} 13 & -4 \\ 5 & -1 \end{bmatrix}$$

you found in Preview Activity 8.1 using technology?

We now consider this result in the general case of a $k \times m$ matrix A and an $m \times n$ matrix B , where A and B define matrix transformations S and T , respectively. In other words, S and T are matrix transformations defined by $S(\mathbf{x}) = A\mathbf{x}$ and $T(\mathbf{x}) = B\mathbf{x}$. The domain of S is \mathbb{R}^m and the codomain is \mathbb{R}^k . The domain of T is \mathbb{R}^n and the codomain is \mathbb{R}^m . The composition $S \circ T$ is defined because the output vectors of T are in \mathbb{R}^m and they lie in the domain of S . The domain of $S \circ T$ is the same as the domain of T since the input vectors first go through the T transformation. The codomain of $S \circ T$ is the same as the codomain of S since the final output vectors are produced by applying the S transformation.

Let us see how we can obtain the matrix corresponding to the transformation $S \circ T$. Let $B = [\mathbf{b}_1 \ \mathbf{b}_2 \ \cdots \ \mathbf{b}_n]$, where \mathbf{b}_j is the j th column of B , and let $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$. Recall that the matrix

vector product $B\mathbf{x}$ is the linear combination of the columns of B with the corresponding weights from \mathbf{x} . So

$$T(\mathbf{x}) = B\mathbf{x} = x_1\mathbf{b}_1 + x_2\mathbf{b}_2 + \cdots + x_n\mathbf{b}_n.$$

Note that each of the \mathbf{b}_j vectors are in \mathbb{R}^m since B is an $m \times n$ matrix. Therefore, each of these vectors can be multiplied by matrix A and we can evaluate $S(B\mathbf{x})$. Therefore, $S \circ T$ is defined and

$$(S \circ T)(\mathbf{x}) = S(T(\mathbf{x})) = A(B\mathbf{x}) = A(x_1\mathbf{b}_1 + x_2\mathbf{b}_2 + \cdots + x_n\mathbf{b}_n). \quad (8.1)$$



The properties of matrix-vector products show that

$$A(x_1\mathbf{b}_1 + x_2\mathbf{b}_2 + \cdots + x_n\mathbf{b}_n) = x_1A\mathbf{b}_1 + x_2A\mathbf{b}_2 + \cdots + x_nA\mathbf{b}_n. \quad (8.2)$$

This expression is a linear combination of $A\mathbf{b}_i$'s with x_i 's being the weights. Therefore, if we let C be the matrix with columns $A\mathbf{b}_1, A\mathbf{b}_2, \dots, A\mathbf{b}_n$, that is

$$C = [A\mathbf{b}_1 \ A\mathbf{b}_2 \ \cdots \ A\mathbf{b}_n],$$

then

$$x_1A\mathbf{b}_1 + x_2A\mathbf{b}_2 + \cdots + x_nA\mathbf{b}_n = C\mathbf{x} \quad (8.3)$$

by definition of the matrix-vector product. Combining equations (8.1), (8.2), and (8.3) shows that

$$(S \circ T)(\mathbf{x}) = C\mathbf{x}$$

where $C = [A\mathbf{b}_1 \ A\mathbf{b}_2 \ \cdots \ A\mathbf{b}_n]$.

Also note that since $T(\mathbf{x}) = B\mathbf{x}$ and $S(\mathbf{y}) = A\mathbf{y}$, we find

$$(S \circ T)(\mathbf{x}) = S(T(\mathbf{x})) = S(B\mathbf{x}) = A(B(\mathbf{x})). \quad (8.4)$$

Since the matrix representing the transformation $S \circ T$ is the matrix

$$[A\mathbf{b}_1 \ A\mathbf{b}_2 \ \cdots \ A\mathbf{b}_n]$$

where $\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_n$ are the columns of the matrix B , it is natural to define AB to be this matrix in light of equation (8.4).

Matrix multiplication has some properties that are unfamiliar to us as the next activity illustrates.

Activity 8.2. Let $A = \begin{bmatrix} 3 & -1 \\ -2 & 6 \end{bmatrix}$, $B = \begin{bmatrix} 0 & 2 \\ 1 & 3 \end{bmatrix}$, $C = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$, $D = \begin{bmatrix} 3 & -3 \\ -3 & 3 \end{bmatrix}$ and $E = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$.

- (a) Find the indicated products (by hand or using a calculator).

$$AB \quad BA \quad DC \quad AC \quad BC \quad AE \quad EB$$

- (b) Is matrix multiplication commutative? Explain.

- (c) Is there an identity element for matrix multiplication? In other words, is there a matrix I for which $AI = IA = A$ for any matrix A ? Explain.

- (d) If a and b are real numbers with $ab = 0$, then we know that either $a = 0$ or $b = 0$. Is this same property true with matrix multiplication? Explain.

- (e) If a , b , and c are real numbers with $c \neq 0$ and $ac = bc$, we know that $a = b$. Is this same property true with matrix multiplication? Explain.

As we saw in Activity 8.2, there are matrices A, B for which $AB \neq BA$. On the other hand, there are matrices for which $AB = BA$. For example, this equality will always hold for a square matrix A and if B is the identity matrix of the same size. It also holds if $A = B$. If the equality



$AB = BA$ holds, we say that matrices A and B *commute*. So the identity matrix commutes with all square matrices of the same size and every matrix A commutes with A^k for any power k .

There is an alternative method of calculating a matrix product that we will often use that we illustrate in the next activity. This alternate version depends on the product of a row matrix with a

vector. Suppose $A = [a_1 \ a_2 \ \cdots \ a_n]$ is a $1 \times n$ matrix and $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$ is an $n \times 1$ vector. Then

the product $A\mathbf{x}$ is the 1×1 vector

$$[a_1 \ a_2 \ \cdots \ a_n] \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = [a_1x_1 + a_2x_2 + \cdots + a_nx_n].$$

In this situation, we usually identify the 1×1 matrix with its scalar entry and write

$$[a_1 \ a_2 \ \cdots \ a_n] \cdot \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = a_1x_1 + a_2x_2 + \cdots + a_nx_n. \quad (8.5)$$

The product \cdot in (8.5) is called the *scalar* or *dot* product of $[a_1 \ a_2 \ \cdots \ a_n]$ with $\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$.

Activity 8.3. Let $A = \begin{bmatrix} 1 & -1 & 2 \\ 3 & 0 & -4 \\ 2 & -5 & 1 \end{bmatrix}$ and $B = \begin{bmatrix} 4 & -2 \\ 6 & 0 \\ 1 & 3 \end{bmatrix}$.

Let \mathbf{a}_i be the i th row of A and \mathbf{b}_j the j th column of B . For example, $\mathbf{a}_1 = [1 \ -1 \ 2]$ and $\mathbf{b}_2 = \begin{bmatrix} -2 \\ 0 \\ 3 \end{bmatrix}$.

Calculate the entries of the matrix C , where

$$C = \begin{bmatrix} \mathbf{a}_1 \cdot \mathbf{b}_1 & \mathbf{a}_1 \cdot \mathbf{b}_2 \\ \mathbf{a}_2 \cdot \mathbf{b}_1 & \mathbf{a}_2 \cdot \mathbf{b}_2 \\ \mathbf{a}_3 \cdot \mathbf{b}_1 & \mathbf{a}_3 \cdot \mathbf{b}_2 \end{bmatrix},$$

where $\mathbf{a}_i \cdot \mathbf{b}_j$ refers to the scalar product of row i of A with column j of B .² Compare your result with the result of AB calculated via the product of A with the columns of B .

²Recall from Exercise 5 of Section 5 that the scalar product $\mathbf{u} \cdot \mathbf{v}$ of a $1 \times n$ matrix $\mathbf{u} = [u_1 \ u_2 \ \dots \ u_n]$ and an $n \times 1$ vector $\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$ is $\mathbf{u} \cdot \mathbf{v} = u_1v_1 + u_2v_2 + u_3v_3 + \cdots + u_nv_n$.

Activity 8.3 shows that there is an alternate way to calculate a matrix product. To see how this works in general, let $A = [a_{ij}]$ be a $k \times m$ matrix and $B = [\mathbf{b}_1 \ \mathbf{b}_2 \ \cdots \ \mathbf{b}_n]$ an $m \times n$ matrix. We know that

$$AB = [A\mathbf{b}_1 \ A\mathbf{b}_2 \ \cdots \ A\mathbf{b}_n].$$

Now let $\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_k$ be the rows of A so that $A = \begin{bmatrix} \mathbf{r}_1 \\ \mathbf{r}_2 \\ \vdots \\ \mathbf{r}_k \end{bmatrix}$. First we argue that if $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{bmatrix}$,

then

$$A\mathbf{x} = \begin{bmatrix} \mathbf{r}_1 \cdot \mathbf{x} \\ \mathbf{r}_2 \cdot \mathbf{x} \\ \vdots \\ \mathbf{r}_k \cdot \mathbf{x} \end{bmatrix}.$$

This is the *scalar product* (or *dot product*) definition of the matrix-vector product.

To show that this definition gives the same result as the linear combination definition of matrix-vector product, we first let $A = [\mathbf{c}_1 \ \mathbf{c}_2 \ \cdots \ \mathbf{c}_m]$, where $\mathbf{c}_1, \mathbf{c}_2, \dots, \mathbf{c}_m$ are the columns of A . By our linear combination definition of the matrix-vector product, we obtain

$$\begin{aligned} A\mathbf{x} &= x_1\mathbf{c}_1 + x_2\mathbf{c}_2 + \cdots + x_m\mathbf{c}_m \\ &= x_1 \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{k1} \end{bmatrix} + x_2 \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{k2} \end{bmatrix} + \cdots + x_m \begin{bmatrix} a_{1m} \\ a_{2m} \\ \vdots \\ a_{km} \end{bmatrix} \\ &= \begin{bmatrix} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1m}x_m \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2m}x_m \\ \vdots \\ a_{k1}x_1 + a_{k2}x_2 + \cdots + a_{km}x_m \end{bmatrix} \\ &= \begin{bmatrix} \mathbf{r}_1 \cdot \mathbf{x} \\ \mathbf{r}_2 \cdot \mathbf{x} \\ \vdots \\ \mathbf{r}_k \cdot \mathbf{x} \end{bmatrix}. \end{aligned}$$

Therefore, the above work shows that both linear combination and scalar product definitions give the same matrix-vector product.

Applying this to the matrix product AB defined in terms of the matrix-vector product, we see that

$$A\mathbf{b}_j = \begin{bmatrix} \mathbf{r}_1 \cdot \mathbf{b}_j \\ \mathbf{r}_2 \cdot \mathbf{b}_j \\ \vdots \\ \mathbf{r}_k \cdot \mathbf{b}_j \end{bmatrix}.$$

So the i, j th entry of the matrix product AB is found by taking the scalar product of the i th row of A with the j th column of B . In other words,

$$(AB)_{ij} = \mathbf{r}_i \cdot \mathbf{b}_j$$



where \mathbf{r}_i is the i th row of A and \mathbf{b}_j is the j th column of B .

Properties of Matrix Multiplication

Activity 8.2 shows that we must be very careful not to assume that matrix multiplication behaves like multiplication of real numbers. However, matrix multiplication does satisfy some familiar properties. For example, we now have an addition and multiplication of matrices under certain conditions, so we might ask if matrix multiplication distributes over matrix addition. To answer this question we take two *arbitrary* $k \times m$ matrices A and B and an *arbitrary* $m \times n$ matrix $C = [\mathbf{c}_1 \ \mathbf{c}_2 \ \cdots \ \mathbf{c}_n]$. Then

$$\begin{aligned}(A + B)C &= [(A + B)\mathbf{c}_1 \ (A + B)\mathbf{c}_2 \ \cdots \ (A + B)\mathbf{c}_n] \\ &= [A\mathbf{c}_1 + B\mathbf{c}_1 \ A\mathbf{c}_2 + B\mathbf{c}_2 \ \cdots \ A\mathbf{c}_n + B\mathbf{c}_n] \\ &= [A\mathbf{c}_1 \ A\mathbf{c}_2 \ \cdots \ A\mathbf{c}_n] + [B\mathbf{c}_1 \ B\mathbf{c}_2 \ \cdots \ B\mathbf{c}_n] \\ &= AC + BC.\end{aligned}$$

Similar arguments can be used to show the following properties of matrix multiplication.

Theorem 8.3. *Let A , B , and C be matrices of the appropriate sizes for all sums and products to be defined and let a be a scalar. Then*

- (1) $(AB)C = A(BC)$ (*this property tells us that matrix multiplication is associative*)
- (2) $(A + B)C = AC + BC$ (*this property tells us that matrix multiplication on the right distributes over matrix addition*)
- (3) $A(B+C) = AB+AC$ (*this property tells us that matrix multiplication on the left distributes over matrix addition*)
- (4) *There is a square matrix I_n with the property that $AI_n = A$ or $I_nA = A$ for whichever product is defined.*
- (5) $a(AB) = (aA)B = A(aB)$

We verified the second part of this theorem and will assume that all of the properties of this theorem hold. The matrix I_n introduced in Theorem 8.3 is called the (*multiplicative*) *identity matrix*. We usually omit the word multiplicative and refer to the I_n simply as the identity matrix. This does not cause any confusion since we refer to the additive identity matrix as simply the zero matrix.

Definition 8.4. Let n be a positive integer. The $n \times n$ **identity matrix** I_n is the matrix $I_n = [a_{ij}]$, where $a_{ii} = 1$ for each i and $a_{ij} = 0$ if $i \neq j$.

We also write the matrix I_n as

$$I_n = \begin{bmatrix} 1 & 0 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & 0 & \cdots & 0 & 0 \\ \vdots & & & & \ddots & & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 1 & 0 \\ 0 & 0 & 0 & 0 & \cdots & 0 & 1 \end{bmatrix}.$$



The matrix I_n has the property that for any $n \times n$ matrix A ,

$$AI_n = I_n A = A.$$

so I_n is a multiplicative identity in the set of all $n \times n$ matrices. More generally, for an $m \times n$ matrix A ,

$$AI_n = I_m A = A.$$

The Transpose of a Matrix

One additional operation on matrices is the transpose. The transpose of a matrix occurs in many useful formulas in linear algebra and in applications of linear algebra.

Definition 8.5. The **transpose** of an $m \times n$ matrix $A = [a_{ij}]$ is the $n \times m$ matrix A^T whose i, j th entry is a_{ji} .

Written out, the transpose of the $m \times n$ matrix

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n-1} & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n-1} & a_{2n} \\ \vdots & & \ddots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn-1} & a_{mn} \end{bmatrix}$$

is the $n \times m$ matrix

$$A^T = \begin{bmatrix} a_{11} & a_{21} & \cdots & a_{m-11} & a_{m1} \\ a_{12} & a_{22} & \cdots & a_{m-12} & a_{m2} \\ \vdots & & \ddots & & \vdots \\ a_{1n} & a_{2n} & \cdots & a_{m-1n} & a_{mn} \end{bmatrix}.$$

In other words, the transpose of a matrix A is the matrix A^T whose rows are the columns of A . Alternatively, the transpose of A is the matrix A^T whose columns are the rows of A . We can also view the transpose of A as the reflection of A across its main diagonal, where the *diagonal* of a matrix $A = [a_{ij}]$ consists of the entries of the form $[a_{ii}]$.

Activity 8.4.

- (a) Find the transpose of each of the indicated matrices.

$$\begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \end{bmatrix} \quad \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} \quad \begin{bmatrix} 1 & 2 \\ 4 & -3 \\ 0 & -1 \end{bmatrix}$$

- (b) Find the transpose of the new matrix for each part above. What can you conjecture based on your results?
- (c) There are certain special types of matrices that are given names.

Definition 8.6. Let A be a square matrix whose ij th entry is a_{ij} .

- (1) The matrix A is a **diagonal matrix** if $a_{ij} = 0$ whenever $i \neq j$.



- (2) The matrix A is a **symmetric** matrix if $A^T = A$.
- (3) The matrix A is an **upper triangular** if $a_{ij} = 0$ whenever $i > j$.
- (4) The matrix A is a **lower triangular** if $a_{ij} = 0$ whenever $i < j$.
- Find an example of a diagonal matrix A . What can you say about A^T ?
 - Find an example of a non-diagonal symmetric matrix B . If $B^T = B$, must B be a square matrix?
 - Find an example of an upper triangular matrix C . What kind of a matrix is C^T ?

We will see later that diagonal matrices are important in that their powers are easy to calculate. Symmetric matrices arise frequently in applications such as in graph theory as adjacency matrices and in quantum mechanics as observables, and have many useful properties including being diagonalizable and having real eigenvalues, as we will also see later.

Properties of the Matrix Transpose

As with every other operation, we want to understand what properties the matrix transpose has. Properties of transposes are shown in the following theorem.

Theorem 8.7. *Let A and B be matrices of the appropriate sizes and let a be a scalar. Then*

- (1) $(A^T)^T = A$
- (2) $(A + B)^T = A^T + B^T$
- (3) $(AB)^T = B^T A^T$
- (4) $(aA)^T = aA^T$

The one property that might seem strange is the third one. To understand this property, suppose A is an $m \times n$ matrix and B an $n \times k$ matrix so that the product AB is defined. We will argue that $(AB)^T = B^T A^T$ by comparing the i, j th entry of each side.

- First notice that the i, j th entry of $(AB)^T$ is the j, i th entry of AB . The j, i th entry of AB is found by taking the scalar product of the j th row of A with the i th column of B . Thus,

the i, j th entry of $(AB)^T$ is the scalar product of the j th row of A with the i th column of B .

- The i, j th entry of $B^T A^T$ is the scalar product of the i th row of B^T with the j th column of A^T . But the i th row of B^T is the i th column of B and the j th column of A^T is the j th row of A . So

the i, j th entry of $B^T A^T$ is the scalar product of the j th row of A with the i th column of B .

Since the two matrices $(AB)^T$ and $B^T A^T$ have the same size and same corresponding entries, they are the same matrix.



Examples

What follows are worked examples that use the concepts from this section.

Example 8.8. Let

$$\begin{aligned} A &= \begin{bmatrix} 1 & 2 & 0 & 1 \\ 3 & 0 & -4 & 5 \\ 7 & 6 & -1 & 0 \end{bmatrix} & B &= \begin{bmatrix} -2 & 4 & -3 \\ 5 & 1 & 9 \\ 1 & 1 & -2 \end{bmatrix} \\ C &= \begin{bmatrix} 0 & -1 & 6 \\ 3 & -2 & 5 \\ 1 & 0 & 4 \end{bmatrix} & D &= \begin{bmatrix} 10 & -4 \\ 5 & 2 \\ 8 & -1 \end{bmatrix} \\ E &= \begin{bmatrix} 1 & 0 \\ 4 & -3 \\ 5 & -1 \end{bmatrix} & \text{and} & F = \begin{bmatrix} -2 & 1 & 5 \\ 6 & 3 & -8 \\ 1 & 0 & -1 \\ 7 & 0 & -5 \end{bmatrix}. \end{aligned}$$

Determine the results of the following operations, if defined. If not defined, explain why.

- (a) AF
- (b) $A(BC)$
- (c) $(BC)A$
- (d) $(B + C)D$
- (e) $D^T E$
- (f) $(A^T + F)^T$

Example Solution.

- (a) Since A is a 3×4 matrix and F is a 4×3 matrix, the number of columns of A equals the number of rows of F and the matrix product AF is defined. Recall that if $F = [f_1 \ f_2 \ f_3]$, where f_1, f_2, f_3 are the columns of F , then $AF = [Af_1 \ Af_2 \ Af_3]$. Recall also that Af_1 is the linear combination of the columns of A with weights from f_1 , so

$$\begin{aligned} Af_1 &= \begin{bmatrix} 1 & 2 & 0 & 1 \\ 3 & 0 & -4 & 5 \\ 7 & 6 & -1 & 0 \end{bmatrix} \begin{bmatrix} -2 \\ 6 \\ 1 \\ 7 \end{bmatrix} \\ &= (-2) \begin{bmatrix} 1 \\ 3 \\ 7 \end{bmatrix} + (6) \begin{bmatrix} 2 \\ 0 \\ 6 \end{bmatrix} + (1) \begin{bmatrix} 0 \\ -4 \\ -1 \end{bmatrix} + (7) \begin{bmatrix} 1 \\ 5 \\ 0 \end{bmatrix} \\ &= \begin{bmatrix} -2 + 12 + 0 + 7 \\ -6 + 0 - 4 + 35 \\ -14 + 36 - 1 + 0 \end{bmatrix} \\ &= \begin{bmatrix} 17 \\ 25 \\ 21 \end{bmatrix}, \end{aligned}$$

$$\begin{aligned}
A\mathbf{f}_2 &= \begin{bmatrix} 1 & 2 & 0 & 1 \\ 3 & 0 & -4 & 5 \\ 7 & 6 & -1 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 3 \\ 0 \\ 0 \end{bmatrix} \\
&= (1) \begin{bmatrix} 1 \\ 3 \\ 7 \end{bmatrix} + (3) \begin{bmatrix} 2 \\ 0 \\ 6 \end{bmatrix} + (0) \begin{bmatrix} 0 \\ -4 \\ -1 \end{bmatrix} + (0) \begin{bmatrix} 1 \\ 5 \\ 0 \end{bmatrix} \\
&= \begin{bmatrix} 1+6+0+0 \\ 3+0+0+0 \\ 7+18+0+0 \end{bmatrix} \\
&= \begin{bmatrix} 7 \\ 3 \\ 25 \end{bmatrix},
\end{aligned}$$

and

$$\begin{aligned}
A\mathbf{f}_3 &= \begin{bmatrix} 1 & 2 & 0 & 1 \\ 3 & 0 & -4 & 5 \\ 7 & 6 & -1 & 0 \end{bmatrix} \begin{bmatrix} 5 \\ -8 \\ -1 \\ -5 \end{bmatrix} \\
&= (5) \begin{bmatrix} 1 \\ 3 \\ 7 \end{bmatrix} - (8) \begin{bmatrix} 2 \\ 0 \\ 6 \end{bmatrix} - (1) \begin{bmatrix} 0 \\ -4 \\ -1 \end{bmatrix} - (5) \begin{bmatrix} 1 \\ 5 \\ 0 \end{bmatrix} \\
&= \begin{bmatrix} 5-16-0-5 \\ 15-0+4-25 \\ 35-48+1-0 \end{bmatrix} \\
&= \begin{bmatrix} -16 \\ -6 \\ -12 \end{bmatrix}.
\end{aligned}$$

$$\text{So } AF = \begin{bmatrix} 17 & 7 & -16 \\ 25 & 3 & -6 \\ 21 & 25 & -12 \end{bmatrix}.$$

Alternatively, if $A = \begin{bmatrix} \mathbf{a}_1 \\ \mathbf{a}_2 \\ \mathbf{a}_3 \\ \mathbf{a}_4 \end{bmatrix}$, then the matrix product AF is the matrix whose ij entry is $\mathbf{a}_i \cdot \mathbf{f}_j$. Using this method we have

$$AF = \begin{bmatrix} \mathbf{a}_1 \cdot \mathbf{f}_1 & \mathbf{a}_1 \cdot \mathbf{f}_2 & \mathbf{a}_1 \cdot \mathbf{f}_3 \\ \mathbf{a}_2 \cdot \mathbf{f}_1 & \mathbf{a}_2 \cdot \mathbf{f}_2 & \mathbf{a}_2 \cdot \mathbf{f}_3 \\ \mathbf{a}_3 \cdot \mathbf{f}_1 & \mathbf{a}_3 \cdot \mathbf{f}_2 & \mathbf{a}_3 \cdot \mathbf{f}_3 \end{bmatrix}.$$



Now

$$\begin{aligned}
 \mathbf{a}_1 \cdot \mathbf{f}_1 &= (1)(-2) + (2)(6) + (0)(1) + (1)(7) = 17 \\
 \mathbf{a}_1 \cdot \mathbf{f}_2 &= (1)(1) + (2)(3) + (0)(0) + (1)(0) = 7 \\
 \mathbf{a}_1 \cdot \mathbf{f}_3 &= (1)(5) + (2)(-8) + (0)(-1) + (1)(-5) = -16 \\
 \mathbf{a}_2 \cdot \mathbf{f}_1 &= (3)(-2) + (0)(6) + (-4)(1) + (5)(7) = 25 \\
 \mathbf{a}_2 \cdot \mathbf{f}_2 &= (3)(1) + (0)(3) + (-4)(0) + (5)(0) = 3 \\
 \mathbf{a}_2 \cdot \mathbf{f}_3 &= (3)(5) + (0)(-8) + (-4)(-1) + (5)(-5) = -6 \\
 \mathbf{a}_3 \cdot \mathbf{f}_1 &= (7)(-2) + (6)(6) + (-1)(1) + (0)(7) = 21 \\
 \mathbf{a}_3 \cdot \mathbf{f}_2 &= (7)(1) + (6)(3) + (-1)(0) + (0)(0) = 25 \\
 \mathbf{a}_3 \cdot \mathbf{f}_3 &= (7)(5) + (6)(-8) + (-1)(-1) + (0)(-5) = -12,
 \end{aligned}$$

$$\text{so } AF = \begin{bmatrix} 17 & 7 & -16 \\ 25 & 3 & -6 \\ 21 & 25 & -12 \end{bmatrix}.$$

- (b) Since BC is a 3×3 matrix but A is 3×4 , the number of columns of A is not equal to the number of rows of BC . We conclude that $A(BC)$ is not defined.
- (c) Since BC is a 3×3 matrix and A is 3×4 , the number of columns of BC is equal to the number of rows of A . Thus, the quantity $(BC)A$ is defined. First we calculate BC using the dot product of the rows of B with the columns of C . Letting $B = \begin{bmatrix} \mathbf{b}_1 \\ \mathbf{b}_2 \\ \mathbf{b}_3 \end{bmatrix}$ and $C = [\mathbf{c}_1 \ \mathbf{c}_2 \ \mathbf{c}_3]$, where $\mathbf{b}_1, \mathbf{b}_2$, and \mathbf{b}_3 are the rows of B and $\mathbf{c}_1, \mathbf{c}_2$, and \mathbf{c}_3 are the columns of C , we have

$$BC = \begin{bmatrix} \mathbf{b}_1 \cdot \mathbf{c}_1 & \mathbf{b}_1 \cdot \mathbf{c}_2 & \mathbf{b}_1 \cdot \mathbf{c}_3 \\ \mathbf{b}_2 \cdot \mathbf{c}_1 & \mathbf{b}_2 \cdot \mathbf{c}_2 & \mathbf{b}_2 \cdot \mathbf{c}_3 \\ \mathbf{b}_3 \cdot \mathbf{c}_1 & \mathbf{b}_3 \cdot \mathbf{c}_2 & \mathbf{b}_3 \cdot \mathbf{c}_3 \end{bmatrix}.$$

Now

$$\begin{aligned}
 \mathbf{b}_1 \cdot \mathbf{c}_1 &= (-2)(0) + (4)(3) + (-3)(1) = 9 \\
 \mathbf{b}_1 \cdot \mathbf{c}_1 &= (-2)(-1) + (4)(-2) + (-3)(0) = -6 \\
 \mathbf{b}_1 \cdot \mathbf{c}_1 &= (-2)(6) + (4)(5) + (-3)(4) = -4 \\
 \mathbf{b}_1 \cdot \mathbf{c}_1 &= (5)(0) + (1)(3) + (9)(1) = 12 \\
 \mathbf{b}_1 \cdot \mathbf{c}_1 &= (5)(-1) + (1)(-2) + (9)(0) = -7 \\
 \mathbf{b}_1 \cdot \mathbf{c}_1 &= (5)(6) + (1)(5) + (9)(4) = 71 \\
 \mathbf{b}_1 \cdot \mathbf{c}_1 &= (1)(0) + (1)(3) + (-2)(1) = 1 \\
 \mathbf{b}_1 \cdot \mathbf{c}_1 &= (1)(-1) + (1)(-2) + (-2)(0) = -3 \\
 \mathbf{b}_1 \cdot \mathbf{c}_1 &= (1)(6) + (1)(5) + (-2)(4) = 3,
 \end{aligned}$$

$$\text{so } BC = \begin{bmatrix} 9 & -6 & -4 \\ 12 & -7 & 71 \\ 1 & -3 & 3 \end{bmatrix}. \text{ If } BC = \begin{bmatrix} \mathbf{r}_1 \\ \mathbf{r}_2 \\ \mathbf{r}_3 \end{bmatrix} \text{ and } A = [\mathbf{s}_1 \ \mathbf{s}_2 \ \mathbf{s}_3 \ \mathbf{s}_4], \text{ where } \mathbf{r}_1, \mathbf{r}_2, \text{ and }$$



\mathbf{r}_3 are the rows of BC and $\mathbf{s}_1, \mathbf{s}_2, \mathbf{s}_3$, and \mathbf{s}_4 are the columns of A , then

$$(BC)A = \begin{bmatrix} \mathbf{r}_1 \cdot \mathbf{s}_1 & \mathbf{r}_1 \cdot \mathbf{s}_2 & \mathbf{r}_1 \cdot \mathbf{s}_3 & \mathbf{r}_1 \cdot \mathbf{s}_4 \\ \mathbf{r}_2 \cdot \mathbf{s}_1 & \mathbf{r}_2 \cdot \mathbf{s}_2 & \mathbf{r}_2 \cdot \mathbf{s}_3 & \mathbf{r}_2 \cdot \mathbf{s}_4 \\ \mathbf{r}_3 \cdot \mathbf{s}_1 & \mathbf{r}_3 \cdot \mathbf{s}_2 & \mathbf{r}_3 \cdot \mathbf{s}_3 & \mathbf{r}_3 \cdot \mathbf{s}_4 \end{bmatrix}.$$

Now

$$\begin{aligned} \mathbf{r}_1 \cdot \mathbf{s}_1 &= (9)(1) + (-6)(3) + (-4)(7) = -37 \\ \mathbf{r}_1 \cdot \mathbf{s}_2 &= (9)(2) + (-6)(0) + (-4)(6) = -6 \\ \mathbf{r}_1 \cdot \mathbf{s}_3 &= (9)(0) + (-6)(-4) + (-4)(-1) = 28 \\ \mathbf{r}_1 \cdot \mathbf{s}_4 &= (9)(1) + (-6)(5) + (-4)(0) = -21 \\ \mathbf{r}_2 \cdot \mathbf{s}_1 &= (12)(1) + (-7)(3) + (71)(7) = 488 \\ \mathbf{r}_2 \cdot \mathbf{s}_2 &= (12)(2) + (-7)(0) + (71)(6) = 450 \\ \mathbf{r}_2 \cdot \mathbf{s}_3 &= (12)(0) + (-7)(-4) + (71)(-1) = -43 \\ \mathbf{r}_2 \cdot \mathbf{s}_4 &= (12)(1) + (-7)(5) + (71)(0) = -23 \\ \mathbf{r}_3 \cdot \mathbf{s}_1 &= (1)(1) + (-3)(3) + (3)(7) = 13 \\ \mathbf{r}_3 \cdot \mathbf{s}_2 &= (1)(2) + (-3)(0) + (3)(6) = 20 \\ \mathbf{r}_3 \cdot \mathbf{s}_3 &= (1)(0) + (-3)(-4) + (3)(-1) = 9 \\ \mathbf{r}_3 \cdot \mathbf{s}_4 &= (1)(1) + (-3)(5) + (3)(0) = -14, \end{aligned}$$

$$\text{so } (BC)A = \begin{bmatrix} -37 & -6 & 28 & -21 \\ 488 & 450 & -43 & -23 \\ 13 & 20 & 9 & -14 \end{bmatrix}.$$

- (d) Since B and C are both 3×3 matrices, their sum is defined and is a 3×3 matrix. Because D is 3×2 matrix, the number of columns of $B + C$ is equal to the number of rows of D . Thus, the quantity $(B + C)D$ is defined and, using the row-column method of matrix multiplication as earlier,

$$\begin{aligned} (B + C)D &= \left(\begin{bmatrix} -2 & 4 & -3 \\ 5 & 1 & 9 \\ 1 & 1 & -2 \end{bmatrix} + \begin{bmatrix} 0 & -1 & 6 \\ 3 & -2 & 5 \\ 1 & 0 & 4 \end{bmatrix} \right) \begin{bmatrix} 10 & -4 \\ 5 & 2 \\ 8 & -1 \end{bmatrix} \\ &= \begin{bmatrix} -2 + 0 & 4 - 1 & -3 + 6 \\ 5 + 3 & 1 - 2 & 9 + 5 \\ 1 + 1 & 1 + 0 & -2 + 4 \end{bmatrix} \begin{bmatrix} 10 & -4 \\ 5 & 2 \\ 8 & -1 \end{bmatrix} \\ &= \begin{bmatrix} -2 & 3 & 3 \\ 8 & -1 & 14 \\ 2 & 1 & 2 \end{bmatrix} \begin{bmatrix} 10 & -4 \\ 5 & 2 \\ 8 & -1 \end{bmatrix} \\ &= \begin{bmatrix} 19 & 11 \\ 187 & -48 \\ 41 & -8 \end{bmatrix}. \end{aligned}$$

- (e) Since D^T is a 2×3 matrix and E is 3×2 , the number of columns of D^T is equal to the



number of rows of E . Thus, $D^T E$ is defined and

$$\begin{aligned} D^T E &= \begin{bmatrix} 10 & -4 \\ 5 & 2 \\ 8 & -1 \end{bmatrix}^T \begin{bmatrix} 1 & 0 \\ 4 & -3 \\ 5 & -1 \end{bmatrix} \\ &= \begin{bmatrix} 10 & 5 & 8 \\ -4 & 2 & -1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 4 & -3 \\ 5 & -1 \end{bmatrix} \\ &= \begin{bmatrix} 70 & -23 \\ -1 & -5 \end{bmatrix}. \end{aligned}$$

- (f) The fact that A is a 3×4 matrix means that A^T is a 4×3 matrix. Since F is also a 4×3 matrix, the sum $A^T + F$ is defined. The transpose of any matrix is also defined, so $(A^T + F)^T$ is defined and

$$\begin{aligned} (A^T + F)^T &= \left(\begin{bmatrix} 1 & 2 & 0 & 1 \\ 3 & 0 & -4 & 5 \\ 7 & 6 & -1 & 0 \end{bmatrix}^T + \begin{bmatrix} -2 & 1 & 5 \\ 6 & 3 & -8 \\ 1 & 0 & -1 \\ 7 & 0 & -5 \end{bmatrix} \right)^T \\ &= \left(\begin{bmatrix} 1 & 3 & 7 \\ 2 & 0 & 6 \\ 0 & -4 & -1 \\ 1 & 5 & 0 \end{bmatrix} + \begin{bmatrix} -2 & 1 & 5 \\ 6 & 3 & -8 \\ 1 & 0 & -1 \\ 7 & 0 & -5 \end{bmatrix} \right)^T \\ &= \left(\begin{bmatrix} 1-2 & 3+1 & 7+5 \\ 2+6 & 0+3 & 6-8 \\ 0+1 & -4+0 & -1-1 \\ 1+7 & 5+0 & 0-5 \end{bmatrix} \right)^T \\ &= \left(\begin{bmatrix} -1 & 4 & 12 \\ 8 & 3 & -2 \\ 1 & -4 & -2 \\ 8 & 5 & -5 \end{bmatrix} \right)^T \\ &= \begin{bmatrix} -1 & 8 & 1 & 8 \\ 4 & 3 & -4 & 5 \\ 12 & -2 & -2 & -5 \end{bmatrix}. \end{aligned}$$

Example 8.9. Let $A = \begin{bmatrix} 2 & -1 \\ 7 & -2 \end{bmatrix}$ and $B = \begin{bmatrix} 4 & 6 \\ -3 & 5 \end{bmatrix}$.

- (a) Determine the matrix sum $A + B$. Then use this sum to calculate $(A + B)^2$.
- (b) Now calculate $(A + B)^2$ in a different way. Use the fact that matrix multiplication distributes over matrix addition to expand (like foiling) $(A + B)^2$ into a sum of matrix products. Then calculate each summand and add to find $(A + B)^2$. You should obtain the same result as part (a). If not, what could be wrong?

Example Solution.



- (a) Adding corresponding terms shows that $A + B = \begin{bmatrix} 6 & 5 \\ 4 & 3 \end{bmatrix}$. Squaring this sum yields the result $(A + B)^2 = \begin{bmatrix} 56 & 45 \\ 36 & 29 \end{bmatrix}$.

- (b) Expanding $(A + B)^2$ (remember that matrix multiplication is not commutative) gives us

$$\begin{aligned}(A + B)^2 &= (A + B)(A + B) \\&= A^2 + AB + BA + B^2 \\&= \begin{bmatrix} -3 & 0 \\ 0 & -3 \end{bmatrix} + \begin{bmatrix} 11 & 7 \\ 34 & 32 \end{bmatrix} + \begin{bmatrix} 50 & -16 \\ 29 & -7 \end{bmatrix} + \begin{bmatrix} -2 & 54 \\ -27 & 7 \end{bmatrix} \\&= \begin{bmatrix} 56 & 45 \\ 36 & 29 \end{bmatrix}\end{aligned}$$

just as in part (a). If instead you obtained the matrix $\begin{bmatrix} 17 & 68 \\ 41 & 68 \end{bmatrix}$ you likely made the mistake of equating $(A + B)^2$ with $A^2 + 2AB + B^2$. These two matrices are not equal in general, because we cannot say that AB is equal to BA .

Summary

In this section we defined a matrix sum, scalar multiples of matrices, the matrix product, and the transpose of a matrix.

- The sum of two $m \times n$ matrices $A = [a_{ij}]$ and $B = [b_{ij}]$ is the $m \times n$ matrix $A + B$ whose i, j th entry is $a_{ij} + b_{ij}$.
- If $A = [a_{ij}]$ is an $m \times n$ matrix, the scalar multiple kA of A by the scalar k is the $m \times n$ matrix whose i, j th entry is ka_{ij} .
- If A is a $k \times m$ matrix and $B = [\mathbf{b}_1 \ \mathbf{b}_2 \ \cdots \ \mathbf{b}_n]$ is an $m \times n$ matrix, then the matrix product AB of the matrices A and B is the $k \times n$ matrix

$$[A\mathbf{b}_1 \ A\mathbf{b}_2 \ \cdots \ A\mathbf{b}_n].$$

The matrix product is defined in this way so that the matrix of a composite $S \circ T$ of linear transformations is the product of matrices of S and T .

- An alternate way of calculating the product of an $k \times m$ matrix A with rows $\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_k$ and an $m \times n$ matrix B with columns $\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_n$ is that the product AB is the $k \times n$ matrix whose i, j th entry is $\mathbf{r}_i \cdot \mathbf{b}_j$.
- Matrix multiplication does not behave as the standard multiplication on real numbers. For example, we can have a product of two non-zero matrices equal to the zero matrix and there is no cancellation law for matrix multiplication.
- The transpose of an $m \times n$ matrix $A = [a_{ij}]$ is the $n \times m$ matrix A^T whose i, j th entry is a_{ji} .



Exercises

- (1) Calculate AB for each of the following matrix pairs by hand in two ways.

$$(a) A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}, B = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

$$(b) A = \begin{bmatrix} 1 & 0 & -1 \end{bmatrix}, B = \begin{bmatrix} 1 & 2 \\ 2 & 3 \\ 3 & 4 \end{bmatrix}$$

- (2) For each of the following A matrices, find all 2×2 matrices $B = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ which commute with the given A . (Two matrices A and B commute with each other if $AB = BA$.)

$$(a) A = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} \quad (b) A = \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix} \quad (c) A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$$

- (3) Find all possible, if any, X matrices satisfying each of the following matrix equations.

$$(a) \begin{bmatrix} 1 & 2 \\ 0 & 2 \end{bmatrix} X = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$$

$$(b) \begin{bmatrix} 1 & -2 \\ -2 & 4 \end{bmatrix} X = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$$

$$(c) \begin{bmatrix} 1 & -2 \\ -2 & 4 \end{bmatrix} X = \begin{bmatrix} 0 & 1 \\ 0 & -2 \end{bmatrix}$$

- (4) For each of the following A matrices, compute $A^2 = AA$, $A^3 = AAA$, A^4 . Use your results to conjecture a formula for A^m . Interpret your answer geometrically using the transformation interpretation.

$$(a) A = \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix} \quad (b) A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \quad (c) A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$$

- (5) If $A\mathbf{v} = 2\mathbf{v}$ for unknown A matrix and \mathbf{v} vector, determine an expression for $A^2\mathbf{v}$, $A^3\mathbf{v}$, ..., $A^m\mathbf{v}$.

- (6) If $A\mathbf{v} = 2\mathbf{v}$ and $A\mathbf{u} = 3\mathbf{u}$, find an expression for $A^m(a\mathbf{v} + b\mathbf{u})$ in terms of \mathbf{v} and \mathbf{u} .

- (7) A matrix A is a **nilpotent** matrix if $A^m = 0$, i.e., A^m is the zero matrix, for some positive integer m . Explain why the matrices

$$A = \begin{bmatrix} 0 & a \\ 0 & 0 \end{bmatrix}, B = \begin{bmatrix} 0 & a & b \\ 0 & 0 & c \\ 0 & 0 & 0 \end{bmatrix}$$

are nilpotent matrices.

- (8) Suppose A is an $n \times n$ matrix for which $A^2 = 0$. Show that there is a matrix B for which $(I_n + A)B = I_n$ where I_n is the identity matrix of size n .



(9) Let A , B , and C be $m \times n$ matrices and let a and b be scalars. Verify Theorem 8.1. That is, show that

- (a) $A + B = B + A$
- (b) $(A + B) + C = A + (B + C)$
- (c) The $m \times n$ matrix 0 whose entries are all 0 has the property that $A + 0 = A$.
- (d) The scalar multiple $(-1)A$ of the matrix A has the property that $(-1)A + A = 0$.
- (e) $(a + b)A = aA + bA$
- (f) $a(A + B) = aA + aB$
- (g) $(ab)A = a(bA)$
- (h) $1A = A$.

(10) Let A , B , and C be matrices of the appropriate sizes for all sums and products to be defined and let a be a scalar. Verify the remaining parts of Theorem 8.3. That is, show that

- (a) $(AB)C = A(BC)$
- (b) $A(B + C) = AB + AC$
- (c) There is a square matrix I_n with the property that $AI_n = A$ or $I_nA = A$ for whichever product is defined.
- (d) $a(AB) = (aA)B = A(aB)$

(11) Let $A = [a_{ij}]$ and $B = [b_{ij}]$ be matrices of the appropriate sizes, and let a be a scalar. Verify the remaining parts of Theorem 8.7. That is, show that

- (a) $(A^T)^T = A$
- (b) $(A + B)^T = A^T + B^T$
- (c) $(aA)^T = aA^T$

(12) The *matrix exponential* is an important tool in solving differential equations. Recall from calculus that the Taylor series expansion for e^x centered at $x = 0$ is

$$e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!} = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots,$$

and that this Taylor series converges to e^x for every real number x . We extend this idea to define the matrix exponential e^A for any square matrix A with real entries as

$$e^A = \sum_{n=0}^{\infty} \frac{1}{n!} A^n = I_n + A + \frac{1}{2!} A^2 + \frac{1}{3!} A^3 + \cdots$$

We explore this idea with an example. Let $B = \begin{bmatrix} 2 & 0 \\ 0 & -1 \end{bmatrix}$.



(a) Calculate B^2, B^3, B^4 . Explain why $B^n = \begin{bmatrix} 2^n & 0 \\ 0 & (-1)^n \end{bmatrix}$ for any positive integer n .

(b) Show that $I_2 + B + B^2 + B^3 + B^4$ is equal to

$$\begin{bmatrix} 1 + 2 + \frac{2^2}{2} + \frac{2^3}{3!} + \frac{2^4}{4!} & 0 \\ 0 & 1 + (-1) + \frac{(-1)^2}{2} + \frac{(-1)^3}{3!} + \frac{(-1)^4}{4!} \end{bmatrix}.$$

(c) Explain why $e^B = \begin{bmatrix} e^2 & 0 \\ 0 & e^{-1} \end{bmatrix}$.

(13) Show that if A and B are 2×2 rotation matrices, then AB is also a 2×2 rotation matrix.

(14) Label each of the following statements as True or False. Provide justification for your response. Throughout, assume that matrices are of the appropriate sizes so that any matrix sums or products are defined.

- (a) **True/False** For any three matrices A, B, C with $A \neq 0$, $AB = AC$ implies $B = C$.
- (b) **True/False** For any three matrices A, B, C with $A \neq 0$, $AB = CA$ implies $B = C$.
- (c) **True/False** If A^2 is the zero matrix, then A itself is the zero matrix.
- (d) **True/False** If $AB = BA$ for every $n \times n$ matrix B , then A is the identity matrix I_n .
- (e) **True/False** If matrix products AB and BA are both defined, then A and B are both square matrices of the same size.
- (f) **True/False** If \mathbf{x}_1 is a solution for $A\mathbf{x} = \mathbf{b}_1$ (i.e., that $A\mathbf{x}_1 = \mathbf{b}_1$) and \mathbf{x}_2 is a solution for $B\mathbf{x} = \mathbf{b}_2$, then $\mathbf{x}_1 + \mathbf{x}_2$ is a solution for $(A + B)\mathbf{x} = \mathbf{b}_1 + \mathbf{b}_2$.
- (g) **True/False** If B is an $m \times n$ matrix with two equal columns, then the matrix AB has two equal columns for every $k \times m$ matrix.
- (h) **True/False** If $A^2 = I_2$, then $A = -I_2$ or $A = I_2$.

Project: Strassen's Algorithm and Partitioned Matrices

Strassen's algorithm is an algorithm for matrix multiplication that can be more efficient than the standard row-column method. To understand this method, we begin with the 2×2 case which will highlight the essential ideas.

Project Activity 8.1. We first work with the 2×2 case.

(a) Let $A = [a_{ij}] = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$ and $B = [b_{ij}] = \begin{bmatrix} 5 & 6 \\ 7 & 8 \end{bmatrix}$.

- i. Calculate the matrix product AB .



- ii. Rather than using eight multiplications to calculate AB , Strassen came up with the idea of using the following seven products:

$$\begin{aligned} h_1 &= (a_{11} + a_{22})(b_{11} + b_{22}) \\ h_2 &= (a_{21} + a_{22})b_{11} \\ h_3 &= a_{11}(b_{12} - b_{22}) \\ h_4 &= a_{22}(b_{21} - b_{11}) \\ h_5 &= (a_{11} + a_{12})b_{22} \\ h_6 &= (a_{21} - a_{11})(b_{11} + b_{12}) \\ h_7 &= (a_{12} - a_{22})(b_{21} + b_{22}). \end{aligned}$$

Calculate h_1 through h_7 for the given matrices A and B . Then calculate the quantities

$$h_1 + h_4 - h_5 + h_7, \quad h_3 + h_5, \quad h_2 + h_4, \quad \text{and } h_1 + h_3 - h_2 + h_6.$$

What do you notice?

- (b) Now we repeat part (a) in general. Suppose we want to calculate the matrix product AB for arbitrary 2×2 matrices $A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$ and $B = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix}$.

Let

$$\begin{aligned} h_1 &= (a_{11} + a_{22})(b_{11} + b_{22}) \\ h_2 &= (a_{21} + a_{22})b_{11} \\ h_3 &= a_{11}(b_{12} - b_{22}) \\ h_4 &= a_{22}(b_{21} - b_{11}) \\ h_5 &= (a_{11} + a_{12})b_{22} \\ h_6 &= (a_{21} - a_{11})(b_{11} + b_{12}) \\ h_7 &= (a_{12} - a_{22})(b_{21} + b_{22}). \end{aligned}$$

Show that

$$AB = \begin{bmatrix} h_1 + h_4 - h_5 + h_7 & h_3 + h_5 \\ h_2 + h_4 & h_1 + h_3 - h_2 + h_6 \end{bmatrix}.$$

The next step is to understand how Strassen's algorithm can be applied to larger matrices. This involves the idea of partitioned (or block) matrices. Recall that the matrix-matrix product of the $k \times m$ matrix A and the $m \times n$ matrix $B = [\mathbf{b}_1 \ \mathbf{b}_2 \ \cdots \ \mathbf{b}_n]$ is defined as

$$AB = [A\mathbf{b}_1 \ A\mathbf{b}_2 \ \cdots \ A\mathbf{b}_n].$$

In this process, we think of B as being partitioned into n columns. We can expand on this idea to partition both A and B when calculating a matrix-matrix product.

Project Activity 8.2. We illustrate the idea of partitioned matrices with an example. Let $A = \begin{bmatrix} 1 & -2 & 3 & -6 & 4 \\ 7 & 5 & 2 & -1 & 0 \\ 3 & -8 & 1 & 0 & 9 \end{bmatrix}$. We can partition A into smaller matrices

$$A = \left[\begin{array}{ccc|cc} 1 & -2 & 3 & -6 & 4 \\ 7 & 5 & 2 & -1 & 0 \\ \hline 3 & -8 & 1 & 0 & 9 \end{array} \right],$$



which are indicated by the vertical and horizontal lines. As a shorthand, we can describe this partition of A as

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix},$$

where $A_{11} = \begin{bmatrix} 1 & -2 & 3 \\ 7 & 5 & 2 \end{bmatrix}$, $A_{12} = \begin{bmatrix} -6 & 4 \\ -1 & 0 \end{bmatrix}$, $A_{21} = [3 \ -8 \ 1]$, and $A_{22} = [0 \ 9]$. The submatrices A_{ij} are called *blocks*. If B is a matrix such that AB is defined, then B must have five

$$\begin{bmatrix} 1 & 3 \\ 2 & 0 \\ 4 & 1 \\ 6 & 5 \\ 4 & 2 \end{bmatrix}$$

rows. As an example, AB is defined if $B = \begin{bmatrix} 1 & 3 \\ 2 & 0 \\ 4 & 1 \\ 6 & 5 \\ 4 & 2 \end{bmatrix}$. The partition of A breaks A up into blocks

with three and two columns, respectively. So if we partition B into blocks with three and two rows, then we can use the blocks to calculate the matrix product AB . For example, partition B as

$$B = \begin{bmatrix} 1 & 3 \\ 2 & 0 \\ 4 & 1 \\ \hline 6 & 5 \\ 4 & 2 \end{bmatrix} = \begin{bmatrix} B_{11} \\ B_{21} \end{bmatrix}.$$

Show that

$$AB = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} B_{11} \\ B_{21} \end{bmatrix} = \begin{bmatrix} A_{11}B_{11} + A_{12}B_{21} \\ A_{21}B_{11} + A_{22}B_{21} \end{bmatrix}.$$

An advantage to using partitioned matrices is that computations with them can be done in parallel, which lessens the time it takes to do the work. In general, we can multiply partitioned matrices as though the submatrices are scalars. That is,

$$\begin{bmatrix} A_{11} & A_{12} & \cdots & A_{1m} \\ A_{21} & A_{22} & \cdots & A_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ A_{i1} & A_{i2} & \cdots & A_{im} \\ \vdots & \vdots & \ddots & \vdots \\ A_{k1} & A_{k2} & \cdots & A_{km} \end{bmatrix} \begin{bmatrix} B_{11} & B_{12} & \cdots & B_{1j} & \cdots & B_{1n} \\ B_{21} & B_{22} & \cdots & B_{2j} & \cdots & B_{2n} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ B_{m1} & B_{m2} & \cdots & B_{mj} & \cdots & B_{mn} \end{bmatrix} = [P_{ij}],$$

where

$$P_{ij} = A_{i1}B_{1j} + A_{i2}B_{2j} + \cdots + A_{im}B_{mj} = \sum_{t=1}^m A_{it}B_{tj},$$

provided that all the submatrix products are defined.

Now we can apply Strassen's algorithm to larger matrices using partitions. This method is sometimes referred to as divide and conquer.

Project Activity 8.3. Let A and B be two $r \times r$ matrices. If r is not a power of 2, then pad the rows and columns of A and B with zeros to make them of size $2^m \times 2^m$ for some integer m . (From a



practical perspective, we might instead just use unequal block sizes.) Let $n = 2^m$. Partition A and B as

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \text{ and } B = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix},$$

where each submatrix is of size $\frac{n}{2} \times \frac{n}{2}$. Now we use the Strassen algorithm just as in the 2×2 case, treating the submatrices as if they were scalars (with the additional constraints of making sure that the dimensions match up so that products are defined, and ensuring we multiply in the correct order). Letting

$$\begin{aligned} M_1 &= (A_{11} + A_{22})(B_{11} + B_{22}) \\ M_2 &= (A_{21} + A_{22})B_{11} \\ M_3 &= A_{11}(B_{12} - B_{22}) \\ M_4 &= A_{22}(B_{21} - B_{11}) \\ M_5 &= (A_{11} + A_{12})B_{22} \\ M_6 &= (A_{21} - A_{11})(B_{11} + B_{12}) \\ M_7 &= (A_{12} - A_{22})(B_{21} + B_{22}), \end{aligned}$$

then the same algebra as in Project Activity 8.1 shows that

$$AB = \begin{bmatrix} M_1 + M_4 - M_5 + M_7 & M_3 + M_5 \\ M_2 + M_4 & M_1 + M_3 - M_2 + M_6 \end{bmatrix}.$$

Apply Strassen's algorithm to calculate the matrix product AB , where

$$A = \begin{bmatrix} 1 & 3 & -1 \\ 2 & 4 & 6 \\ 7 & -2 & 5 \end{bmatrix} \text{ and } B = \begin{bmatrix} 2 & 5 & 3 \\ 2 & -4 & 1 \\ 1 & 6 & 4 \end{bmatrix}.$$

While Strassen's algorithm can be more efficient, it does not always speed up the process. We investigate this in the next activity.

Project Activity 8.4. We introduce a little notation to help us describe the efficiency of our calculations. We won't be formal with this notation, rather work with it in an informal way. Big O (the letter “O”) notation is used to describe the complexity of an algorithm. Generally speaking, in computer science big O notation can be used to describe the run time of an algorithm, the space used by the algorithm, or the number of computations required. The letter “O” is used because the behavior described is also called the order. Big O measures the asymptotic time of an algorithm, not its exact time. For example, if it takes $6n^2 - n + 8$ steps to complete an algorithm, then we say that the algorithm grows at the order of n^2 (we ignore the constants and the smaller power terms, since they become insignificant as n increases) and we describe its growth as $O(n^2)$. To measure the efficiency of an algorithm to determine a matrix product, we will measure the number of operations it takes to calculate the product.

- (a) Suppose A and B are $n \times n$ matrices. Explain why the operation of addition (that is, calculating $A + B$) is $O(n^2)$.
- (b) Suppose A and B are $n \times n$ matrices. How many multiplications are required to calculate the matrix product AB ? Explain.



- (c) The standard algorithm for calculating a matrix product of two $n \times n$ matrices requires n^3 multiplications and a number of additions. Since additions are much less costly in terms of operations, the standard matrix product is $O(n^3)$. We won't show it here, but using Strassen's algorithm on a product of $2^m \times 2^m$ matrices is $O(n^{\log_2(7)})$, where $n = 2^m$. That means that Strassen's algorithm applied to an $n \times n$ matrix (where n is a power of 2) requires approximately $n^{\log_2(7)}$ multiplications. We use this to analyze situations to determine when Strassen's algorithm is computationally more efficient than the standard algorithm.
- i. Suppose A and B are 5×5 matrices. Determine the number of multiplications required to calculate the matrix product AB using the standard matrix product. Then determine the approximate number of multiplications required to calculate the matrix product AB using Strassen's algorithm. Which is more efficient? (Remember, we can only apply Strassen's algorithm to square matrices whose sizes are powers of 2.)
 - ii. Repeat part i. with 125×125 matrices. Which method is more efficient?

As a final note, Strassen's algorithm is approximately $O(n^{2.81})$. As of 2018, the best algorithm for matrix multiplication, developed by Virginia Williams at Stanford University, is approximately $O(n^{2.373})$.³

³V. V. Williams, Multiplying matrices in $O(n^{2.373})$ time, Stanford University, (2014).



Section 9

Introduction to Eigenvalues and Eigenvectors

Focus Questions

By the end of this section, you should be able to give precise and thorough answers to the questions listed below. You may want to keep these questions in mind to focus your thoughts as you complete the section.

- What is an eigenvalue of a matrix?
- What is an eigenvector of a matrix?
- How do we find eigenvectors of a matrix corresponding to an eigenvalue?
- How can the action of a matrix on an eigenvector be visualized?
- Why do we study eigenvalues and eigenvectors?
- What are discrete dynamical systems and how do we analyze the long-term behavior in them?

Application: The Google PageRank Algorithm

The World Wide Web is a vast collection of information, searchable via search engines. A search engine looks for pages that are of interest to the user. In order to be effective, a search engine needs to be able to identify those pages that are relevant to the search criteria provided by the user. This involves determining the relative importance of different web pages by ranking the results of thousands or millions of pages fitting the search criteria. For Google, the PageRank algorithm is their method and is “the heart of our software” as they say. It is this PageRank algorithm that we will learn about later in this section. Eigenvalues and eigenvectors play an important role in this algorithm.

Introduction

Given a matrix A , for some special non-zero vectors \mathbf{v} the action of A on \mathbf{v} will be same as scalar multiplication, i.e., $A\mathbf{v} = \lambda\mathbf{v}$ for some scalar λ . Geometrically, this means that the transformation T defined by $T(\mathbf{x}) = A\mathbf{x}$ simply stretches or contracts the vector \mathbf{v} but does not change its direction. Such a nonzero vector is called an *eigenvector* of A , while the scalar λ is called the corresponding *eigenvalue* of A . The eigenvectors of a matrix tell us quite a bit about the transformation the matrix defines.

Eigenvalues and eigenvectors are used in many applications. Social media like Facebook and Google use eigenvalues to determine the influence of individual members on the network (which can affect advertising) or to rank the importance of web pages. Eigenvalues and eigenvectors appear in quantum physics, where atomic and molecular orbitals can be defined by the eigenvectors of a certain operator. They appear in principal component analysis, used to study large data sets, to diagonalize certain matrices and determine the long term behavior of systems as a result, and in the important singular value decomposition of a matrix. Matrices with real entries can have real or complex eigenvalues, and complex eigenvalues reveal a rotation that is encoded in every real matrix with complex eigenvalues which allows us to better understand certain matrix transformations.

Definition 9.1. Let A be an $n \times n$ matrix. A non-zero vector \mathbf{x} is an **eigenvector** (or **characteristic vector**) of A if there is a scalar λ such that $A\mathbf{x} = \lambda\mathbf{x}$. The scalar λ is an **eigenvalue** (or **characteristic value**) of A .

For example, $\mathbf{v} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ is an eigenvector of $A = \begin{bmatrix} 2 & 1 \\ 3 & 0 \end{bmatrix}$ corresponding to the eigenvalue $\lambda = 3$ because $A\mathbf{v} = \begin{bmatrix} 3 \\ 3 \end{bmatrix}$, which is equal to $3\mathbf{v}$. On the other hand, $\mathbf{w} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ is not an eigenvector of $A = \begin{bmatrix} 2 & 1 \\ 3 & 0 \end{bmatrix}$ because $A\mathbf{w} = \begin{bmatrix} 4 \\ 3 \end{bmatrix}$, which is not a multiple of \mathbf{w} .

Preview Activity 9.1.

- (1) For each of the following parts, use the definition of an eigenvector to determine whether the given vector \mathbf{v} is an eigenvector for the given matrix A . If it is, determine the corresponding eigenvalue.

(a) $A = \begin{bmatrix} 3 & 2 \\ 3 & 8 \end{bmatrix}, \mathbf{v} = \begin{bmatrix} -2 \\ 1 \end{bmatrix}$

(b) $A = \begin{bmatrix} 2 & 0 \\ 0 & -3 \end{bmatrix}, \mathbf{v} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$

(c) $A = \begin{bmatrix} 3 & 2 \\ 3 & 8 \end{bmatrix}, \mathbf{v} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$

(d) $A = \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}, \mathbf{v} = \begin{bmatrix} -2 \\ 1 \end{bmatrix}$

- (2) We now consider how we can find the eigenvectors corresponding to an eigenvalue using the definition. Suppose $A = \begin{bmatrix} 6 & -2 \\ 2 & 1 \end{bmatrix}$. We consider whether we can find eigenvectors



corresponding to eigenvalues 3, and 5. Effectively, this will help us determine whether 3 and/or 5 are eigenvalues of A .

- (a) Rewrite the vector equation $A\mathbf{v} = 5\mathbf{v}$ where $\mathbf{v} = \begin{bmatrix} x \\ y \end{bmatrix}$ as a vector equation.
- (b) After writing $5\mathbf{v}$ as $5I\mathbf{v}$ where I is the identity matrix, rearrange the variables to turn this vector equation into the homogeneous matrix equation $B\mathbf{v} = \mathbf{0}$ where $B = \begin{bmatrix} 1 & -2 \\ 2 & -4 \end{bmatrix}$. If possible, find a non-zero (i.e. a non-trivial) solution to $B\mathbf{v} = \mathbf{0}$. Explain what this means about 5 being an eigenvalue of A or not.
- (c) Similarly, determine whether the vector equation $A\mathbf{v} = 3\mathbf{v}$ has non-zero solutions. Using your result, determine whether 3 is an eigenvalue of A or not.

Eigenvalues and Eigenvectors

Eigenvectors are especially useful in understanding the long-term behavior of dynamical systems, an example of which we will see shortly. The long-term behavior of a dynamical system is quite simple when the initial state vector is an eigenvector and this fact helps us analyze the system in general.

To find eigenvectors, we are interested in determining the vectors \mathbf{x} for which $A\mathbf{x}$ has the same direction as \mathbf{x} . This will happen when

$$A\mathbf{x} = \lambda\mathbf{x}$$

for some scalar λ . Of course, $A\mathbf{x} = \lambda\mathbf{x}$ when $\mathbf{x} = \mathbf{0}$ for every A and every λ , but that is uninteresting. So we really want to consider when there is a non-zero vector \mathbf{x} so that $A\mathbf{x} = \lambda\mathbf{x}$. This prompts the definition of eigenvectors and eigenvalues as in Definition 9.1

In order for a matrix A to have an eigenvector, one condition A must satisfy is that A has to be a square matrix, i.e. an $n \times n$ matrix. We will find that each $n \times n$ matrix has only finitely many eigenvalues.

The terms eigenvalue and eigenvector seem to come from Hilbert, using the German “eigen” (roughly translated as “own”, “proper”, or “characteristic”) to emphasize how eigenvectors and eigenvalues are connected to their matrices. To find the eigenvalues and eigenvectors of an $n \times n$ matrix A , we need to find the solutions to the equation

$$A\mathbf{x} = \lambda\mathbf{x}.$$

In Preview Activity 9.1, we considered this equation for $A = \begin{bmatrix} 6 & -2 \\ 2 & 1 \end{bmatrix}$ and $\lambda = 5$. The homogeneous matrix equation we came up with was

$$\begin{bmatrix} 1 & -2 \\ 2 & -4 \end{bmatrix}\mathbf{x} = \mathbf{0}.$$

To see the relationship between this homogeneous matrix equation and the eigenvalue-eigenvector



equation better, let us consider the eigenvector equation using matrix algebra:

$$\begin{aligned} A\mathbf{x} &= \lambda\mathbf{x} \\ A\mathbf{x} - \lambda\mathbf{x} &= \mathbf{0} \\ A\mathbf{x} - \lambda I_n \mathbf{x} &= \mathbf{0} \\ (A - \lambda I_n)\mathbf{x} &= \mathbf{0}, \end{aligned}$$

where I_n is the $n \times n$ identity matrix. Notice that this description matches the homogenous equation matrix example above since we simply subtracted 5 from the diagonal terms of the matrix A . Hence, to find eigenvalues, we need to find the values of λ so that the homogeneous equation $(A - \lambda I_n)\mathbf{x} = \mathbf{0}$ has non-trivial solutions.

Activity 9.1.

- (a) Under what conditions on $A - \lambda I_n$ will the matrix equation $(A - \lambda I_n)\mathbf{x} = \mathbf{0}$ have non-trivial solutions? Describe at least two different but equivalent conditions.
- (b) The real number 0 is an eigenvalue of $A = \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}$. Check that your criteria in the previous part agrees with this result.
- (c) Determine if 5 is an eigenvalue of the matrix $A = \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}$ using your criterion above.
- (d) What are the two eigenvalues of the matrix $A = \begin{bmatrix} 3 & 2 \\ 4 & 5 \end{bmatrix}$?

Since an eigenvector of A corresponding to eigenvalue λ is a non-trivial solution to the homogeneous equation $(A - \lambda I_n)\mathbf{x} = \mathbf{0}$, the eigenvalues λ which work are those for which the matrix $A - \lambda I_n$ has linearly dependent columns, or for which the row echelon form of the matrix $A - \lambda I_n$ does not have a pivot in every column. When we need to test if a specific λ is an eigenvalue, this method works fine. However, finding which λ 's will work in general involves row reducing a matrix with λ 's subtracted on the diagonal algebraically. For certain types of matrices, this method still provides us the eigenvalues quickly. For general matrices though, row reducing algebraically is not efficient. We will later see an algebraic method which uses the determinants to find the eigenvalues.

Activity 9.2.

- (a) For λ to be an eigenvalue of A , we noted that $A - \lambda I_n$ must have a non-pivot column. Use this criterion to explain why $A = \begin{bmatrix} -2 & 2 \\ 0 & 4 \end{bmatrix}$ has eigenvalues $\lambda = -2$ and $\lambda = 4$.
- (b) Determine the eigenvalues of $A = \begin{bmatrix} 3 & 0 & 1 & 0 \\ 0 & 2 & -1 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$.
- (c) Generalize your results from the above parts in the form of a theorem in the most general $n \times n$ case.



Dynamical Systems

One real-life application of eigenvalues and eigenvectors is in analyzing the long-term behavior of *discrete dynamical systems*. A *dynamical system* is a system of variables whose values change with time. In discrete systems, the change is described by defining the values of the variables at time $t + 1$ in terms of the values at time t . For example, the discrete dynamical system

$$y_{t+1} = y_t + t$$

relates the value of y at time $t + 1$ to the value of y at time t . This is in contrast with a differential equation¹ such as

$$\frac{dy}{dt} = y + t,$$

which describes the instantaneous rate of change of $y(t)$ in terms of y and t .

Discrete dynamical systems can be used in population modeling to provide a simplified model of predator-prey interactions in biology (see Preview Activity 9.2). Other applications include Markov chains (see Exercise 5), age structured population growth models, distillation of a binary ideal mixture of two liquids, cobweb model in economics concerning the interaction of supply and demand for a single good, queuing theory and traffic flow.

Eigenvectors can be used to analyze the long-term behavior of dynamical systems.

Preview Activity 9.2.

- (1) Consider a discrete dynamical system providing a simplified model of predator-prey interactions in biology, such as the system describing the populations of rabbits and foxes in a certain area.

Suppose, for example, for a specific area the model is given by the following equations:

$$\begin{aligned} r_{k+1} &= 1.14r_k - 0.12f_k \\ f_{k+1} &= 0.08r_k + 0.86f_k \end{aligned} \tag{9.1}$$

where r_i represents the number of rabbits in the area i years after a starting time value, and f_i represents the number of foxes in year i . We use r_0, f_0 to denote the initial population values.

- (a) Suppose $r_k = 300$ and $f_k = 100$ for one year. Calculate rabbit and fox population values for the next year. In other words, find r_{k+1}, f_{k+1} values.
- (b) Consider the coefficients of the variables r_k, f_k in the the system of equations in (9.1). Can you explain the reasoning behind the signs and absolute sizes of the coefficients from the story that it models?
- (c) Let $\mathbf{x}_k = \begin{bmatrix} r_k \\ f_k \end{bmatrix}$. The vector \mathbf{x}_k is called the *state vector* of the system at time k , because it describes the state of the whole system at time k . We can rewrite the system of equations in (9.1) as a matrix-vector equation in terms of the state vectors at time k and $k + 1$. More specifically, the equation will be of the form

$$\mathbf{x}_{k+1} = A\mathbf{x}_k \tag{9.2}$$

¹A differential equation is an equation that involves one or more derivatives of a function.



where $A = \begin{bmatrix} 1.14 & -0.12 \\ 0.08 & 0.86 \end{bmatrix}$. We will call this matrix the *transition matrix* of the system. Check that $A \begin{bmatrix} 300 \\ 100 \end{bmatrix}$ gives us the population values you calculated in the first part above.

- (d) The transition matrix will help us simplify calculations of the population values. Note that equation (9.2) implies that $\mathbf{x}_1 = A\mathbf{x}_0$, $\mathbf{x}_2 = A\mathbf{x}_1$, $\mathbf{x}_3 = A\mathbf{x}_2$, and so on. This is a recursive method to find the population values as each year's population values depend on the previous year's population values. Using this approach, calculate \mathbf{x}_k for k values up to 5 corresponding to the following three different initial rabbit-fox population values (all in thousands):

$$r_0 = 300, f_0 = 100$$

$$r_0 = 100, f_0 = 200$$

$$r_0 = 1200, f_0 = 750$$

Can you guess the long-term behavior of the population values in each case? Are they both increasing? Decreasing? One increasing, one decreasing? How do the rabbit and fox populations compare?

A dynamical system is a system of variables whose values change with time. In Preview Activity 9.2, we considered the discrete dynamical system modeling the rabbit and fox population in an area, which is an example of a predator-prey system. The system was given by the equations from (9.1), where r_i represented the number of rabbits in the area i years after a starting time value, and f_i represented the number of foxes in year i . In this notation, r_0, f_0 corresponded to the initial population values.

As we saw in Preview Activity 9.2, if we define the state vector as $\mathbf{x}_k = \begin{bmatrix} r_k \\ f_k \end{bmatrix}$, the system of equations representing the dynamical system can be expressed as

$$\mathbf{x}_{k+1} = A\mathbf{x}_k \tag{9.3}$$

where $A = \begin{bmatrix} 1.14 & -0.12 \\ 0.08 & 0.86 \end{bmatrix}$ represents the transition matrix. Note that equation (9.3) encodes infinitely many equations including $\mathbf{x}_1 = A\mathbf{x}_0$, $\mathbf{x}_2 = A\mathbf{x}_1$, $\mathbf{x}_3 = A\mathbf{x}_2$, and so on. This is a recursive formula for the population values as each year's population values are expressed in terms of the previous year's population values. If we want to calculate \mathbf{x}_{10} , this formula requires first finding the population values for years 1-9. However, we can obtain a non-recursive formula using matrix algebra. If we substitute $\mathbf{x}_1 = A\mathbf{x}_0$ into $\mathbf{x}_2 = A\mathbf{x}_1$ and simplify, we find that

$$\mathbf{x}_2 = A\mathbf{x}_1 = A(A\mathbf{x}_0) = A^2\mathbf{x}_0.$$

Similarly, substituting $\mathbf{x}_2 = A^2\mathbf{x}_0$ into the formula for \mathbf{x}_3 gives

$$\mathbf{x}_3 = A\mathbf{x}_2 = A(A^2\mathbf{x}_0) = A^3\mathbf{x}_0.$$

This process can be continued inductively to show that

$$\mathbf{x}_k = A^k\mathbf{x}_0 \tag{9.4}$$



for every k value. So to find the population values at any year k , we only need to know the initial state vector \mathbf{x}_0 .

Activity 9.3. In this activity the matrix A is the transition matrix for the rabbit and fox population model,

$$A = \begin{bmatrix} 1.14 & -0.12 \\ 0.08 & 0.86 \end{bmatrix}.$$

- (a) Suppose that the initial state vector \mathbf{x}_0 is an eigenvector of A corresponding to eigenvalue λ . In this case, explain why $\mathbf{x}_1 = \lambda\mathbf{x}_0$ and $\mathbf{x}_2 = \lambda^2\mathbf{x}_0$. Find the formula for \mathbf{x}_k in terms of λ , k and \mathbf{x}_0 by applying equation (9.3) iteratively.
- (b) The initial state vector $\mathbf{x}_0 = \begin{bmatrix} 300 \\ 100 \end{bmatrix}$ is an eigenvector of A . Find the corresponding eigenvalue and, using your formula from (a) for \mathbf{x}_k in terms of λ , k and \mathbf{x}_0 , find the state vector \mathbf{x}_k in this case.
- (c) The initial state vector $\mathbf{x}_0 = \begin{bmatrix} 100 \\ 200 \end{bmatrix}$ is an eigenvector of A . Find the corresponding eigenvalue and, using your formula from (a) for \mathbf{x}_k in terms of λ , k and \mathbf{x}_0 , find the state vector \mathbf{x}_k in this case.
- (d) Consider now an initial state vector of the form $\mathbf{x}_0 = a\mathbf{v}_0 + b\mathbf{w}_0$ where a, b are constants, \mathbf{v}_0 is an eigenvector corresponding to eigenvalue λ_1 and \mathbf{w}_0 corresponding to eigenvalue λ_2 (\mathbf{v}_0 and \mathbf{w}_0 are not necessarily the eigenvectors from parts (b) and (c)). Use matrix algebra and equation (9.4) to explain why $\mathbf{x}_k = a\lambda_1^k\mathbf{v}_0 + b\lambda_2^k\mathbf{w}_0$.
- (e) Express the initial state vector $\mathbf{x}_0 = \begin{bmatrix} 1200 \\ 750 \end{bmatrix}$ as a linear combination of the eigenvectors $\mathbf{v}_0 = \begin{bmatrix} 300 \\ 100 \end{bmatrix}$, $\mathbf{w}_0 = \begin{bmatrix} 100 \\ 200 \end{bmatrix}$ and use your result from the previous part to find a formula for \mathbf{x}_k . What happens to the population values as $k \rightarrow \infty$?

As you discovered in Activity 9.3, we can use linearly independent eigenvectors of the transition matrix to find a closed formula for the state vector of a dynamical system, as long as the initial state vector can be expressed as a linear combination of the eigenvectors.

Examples

What follows are worked examples that use the concepts from this section.

Example 9.2. Let $A = \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}$.

- (a) Find all of the eigenvalues of A .
- (b) Find a corresponding eigenvector for each eigenvalue found in part (a).



Example Solution.

- (a) Recall that a scalar λ is a eigenvalue for A if there is a nonzero vector \mathbf{x} such that $A\mathbf{x} = \lambda\mathbf{x}$ or $(A - \lambda I_2)\mathbf{x} = \mathbf{0}$. For this matrix A , we have

$$A - \lambda I_2 = \begin{bmatrix} 1 - \lambda & 2 \\ 2 & 4 - \lambda \end{bmatrix}.$$

To solve the homogeneous system $(A - \lambda I_2)\mathbf{x} = \mathbf{0}$, we row reduce $A - \lambda I_2$. To do this, we first interchange rows to get the following matrix that is row equivalent to $A - \lambda I_2$ (we do this to ensure that we have a nonzero entry in the first row and column)

$$\begin{bmatrix} 2 & 4 - \lambda \\ 1 - \lambda & 2 \end{bmatrix}.$$

Next we replace row two with $\frac{1}{2}(1 - \lambda)$ times row one minus row two to obtain the row equivalent matrix

$$\begin{bmatrix} 2 & 4 - \lambda \\ 0 & \frac{1}{2}(4 - \lambda)(1 - \lambda) - 2 \end{bmatrix}.$$

There will be a nontrivial solution to $(A - \lambda I_2)\mathbf{x} = \mathbf{0}$ if there is a row of zeros in this row echelon form. Thus, we look for values of λ that make

$$\frac{1}{2}(4 - \lambda)(1 - \lambda) - 2 = 0.$$

Applying a little algebra shows that

$$\frac{1}{2}(4 - \lambda)(1 - \lambda) - 2 = 0$$

$$(4 - \lambda)(1 - \lambda) - 4 = 0$$

$$\lambda^2 - 5\lambda = 0$$

$$\lambda(\lambda - 5) = 0.$$

So the eigenvalues of A are $\lambda = 0$ and $\lambda = 5$.

- (b) Recall that an eigenvector for the eigenvalue λ is a nonzero vector \mathbf{x} such that $(A - \lambda I_2)\mathbf{x} = \mathbf{0}$. We consider each eigenvalue in turn.

- When $\lambda = 0$,

$$A - 0I_2 = A = \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}.$$

Technology shows that the reduced row echelon form of A is

$$\begin{bmatrix} 1 & 2 \\ 0 & 0 \end{bmatrix}.$$

If $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$, then $A\mathbf{x} = \mathbf{0}$ implies that x_2 is free and $x_1 = -2x_2$. Choosing

$x_2 = 1$ gives us the eigenvector $\begin{bmatrix} -2 \\ 1 \end{bmatrix}$. As a check, note that

$$\begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix} \begin{bmatrix} -2 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

- When $\lambda = 5$,

$$A - 5I_2 = \begin{bmatrix} -4 & 2 \\ 2 & -1 \end{bmatrix}.$$

Technology shows that the reduced row echelon form of $A - 5I_2$ is

$$\begin{bmatrix} 1 & -\frac{1}{2} \\ 0 & 0 \end{bmatrix}.$$

If $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$, then $(A - 5I_2)\mathbf{x} = \mathbf{0}$ implies that x_2 is free and $x_1 = \frac{1}{2}x_2$. Choosing $x_2 = 2$ gives us the eigenvector $\begin{bmatrix} 1 \\ 2 \end{bmatrix}$. As a check, note that

$$\begin{bmatrix} 1 & -\frac{1}{2} \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

Example 9.3. Accurately predicting the weather has long been an important task. Meteorologists use science, mathematics, and technology to construct models that help us understand weather patterns. These models are very sophisticated, but we will consider only a simple model. Suppose, for example, we want to learn something about whether it will be wet or dry in Grand Rapids, Michigan. To do this, we might begin by collecting some data about weather conditions in Grand Rapids and then use that to make predictions. Information taken over the course of 2017 from the National Weather Service Climate Data shows that if it was dry (meaning no measurable precipitation, either rain or snow) on a given day in Grand Rapids, it would be dry the next day with a probability of 64% and wet with a probability of 36%. Similarly, if it was wet on a given day it would be dry the next day with a probability of 47% and dry with a probability of 53%. Assuming that this pattern is one that continues in the long run, we can develop a mathematical model to make predictions about the weather.

This data tells us how the weather transitions from one day to the next, and we can succinctly represent this data in a *transition matrix*:

$$T = \begin{bmatrix} 0.64 & 0.47 \\ 0.36 & 0.53 \end{bmatrix}. \quad (9.5)$$

Whether it is dry or wet on a given day is called the *state* of that day. So our transition matrix tells us about the transition between states. Notice that if $T = [t_{ij}]$, then the probability of moving from state j to state i is given by t_{ij} . We can represent a state by a vector: the vector $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ represents

the dry state and the vector $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$ represents the wet state.

(a) Calculate $T \begin{bmatrix} 1 \\ 0 \end{bmatrix}$. Interpret the meaning of this output.

(b) Calculate $T \begin{bmatrix} 0 \\ 1 \end{bmatrix}$. Interpret the meaning of this output.

(c) Calculate $T \begin{bmatrix} 0.3 \\ 0.7 \end{bmatrix}$. Interpret the meaning of this output.

(d) We can use the transition matrix to build a chain of probability vectors. We begin with an initial state, say it is dry on a given day. This initial state is represented by the initial state vector $\mathbf{x}_0 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$. The probabilities that it will be dry or wet the following day are given by the vector

$$\mathbf{x}_1 = T\mathbf{x}_0 = \begin{bmatrix} 0.64 \\ 0.36 \end{bmatrix}.$$

This output vector tells us that the next day will be dry with a 64% probability and wet with a 36% probability.

For each $k \geq 1$, we let

$$\mathbf{x}_k = T\mathbf{x}_{k-1}. \quad (9.6)$$

Thus we create a sequence of vectors that tell us the probabilities of it being dry or wet on subsequent days. The vector \mathbf{x}_k is called the *state vector* of the system at time k , because it describes the state of the whole system at time k . We can rewrite the system of equations in (9.1) as a matrix-vector equation in terms of the state vectors at time k and $k + 1$. More specifically, the equation will be of the form

$$\mathbf{x}_{k+1} = T\mathbf{x}_k \quad (9.7)$$

for $k \geq 0$.

- i. Starting with $\mathbf{x}_0 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, use appropriate technology to calculate \mathbf{x}_k for k values up to 10. Round to three decimal places. What do you notice about the entries?
- ii. What does the result of the previous part tell us about eigenvalues of T ? Explain.
- iii. Rewrite T as

$$T = \begin{bmatrix} \frac{64}{100} & \frac{47}{100} \\ \frac{36}{100} & \frac{53}{100} \end{bmatrix}.$$

We do this so we can use exact arithmetic. Let $\mathbf{x} = \begin{bmatrix} \frac{47}{83} \\ \frac{36}{83} \end{bmatrix}$. What is $T\mathbf{x}$? (Use exact arithmetic, no decimals.) Explain how \mathbf{x} is related to the previous two parts of this problem. What does the vector \mathbf{x} tells us about weather in Grand Rapids?

Example Solution.

(a) Here we have

$$T \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0.64 & 0.47 \\ 0.36 & 0.53 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0.64 \\ 0.36 \end{bmatrix}.$$

This output tells us the different probabilities of whether it will be dry or wet the day following a dry day.



(b) Here we have

$$T \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0.64 & 0.47 \\ 0.36 & 0.53 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0.47 \\ 0.53 \end{bmatrix}.$$

This output tells us the different probabilities of whether it will be dry or wet the day following a wet day.

(c) Here we have

$$T \begin{bmatrix} 0.3 \\ 0.7 \end{bmatrix} = \begin{bmatrix} 0.64 & 0.47 \\ 0.36 & 0.53 \end{bmatrix} \begin{bmatrix} 0.3 \\ 0.7 \end{bmatrix} \approx \begin{bmatrix} 0.52 \\ 0.48 \end{bmatrix}.$$

This output tells us there is a 52% chance of it being dry and a 48% chance of it being wet following a day when there is a 30% chance of it being dry and a 70% chance of it being wet.

(d) i. Technology shows that

$$\begin{aligned} \mathbf{x}_1 &= \begin{bmatrix} 0.640 \\ 0.360 \end{bmatrix} & \mathbf{x}_2 &= \begin{bmatrix} 0.579 \\ 0.421 \end{bmatrix} \\ \mathbf{x}_3 &= \begin{bmatrix} 0.568 \\ 0.432 \end{bmatrix} & \mathbf{x}_4 &= \begin{bmatrix} 0.567 \\ 0.433 \end{bmatrix} \\ \mathbf{x}_5 &= \begin{bmatrix} 0.566 \\ 0.434 \end{bmatrix} & \mathbf{x}_6 &= \begin{bmatrix} 0.566 \\ 0.434 \end{bmatrix} \\ \mathbf{x}_7 &= \begin{bmatrix} 0.566 \\ 0.434 \end{bmatrix} & \mathbf{x}_8 &= \begin{bmatrix} 0.566 \\ 0.434 \end{bmatrix} \\ \mathbf{x}_9 &= \begin{bmatrix} 0.566 \\ 0.434 \end{bmatrix} & \mathbf{x}_{10} &= \begin{bmatrix} 0.566 \\ 0.434 \end{bmatrix}. \end{aligned}$$

We can see that our vectors \mathbf{x}_k are essentially the same as we let k increase.

- ii. Since our sequence seems to be converging to a vector \mathbf{x} satisfying $T\mathbf{x} = \mathbf{x}$, we conclude that 1 is an eigenvalue of T .
- iii. A matrix vector multiplication shows that

$$T\mathbf{x} = \begin{bmatrix} \frac{64}{100} & \frac{47}{100} \\ \frac{36}{100} & \frac{53}{100} \end{bmatrix} \begin{bmatrix} \frac{47}{83} \\ \frac{36}{83} \end{bmatrix} = \begin{bmatrix} \frac{47}{83} \\ \frac{36}{83} \end{bmatrix}.$$

In other words, \mathbf{x} is an eigenvector for T with eigenvalue 1. Notice that

$$\frac{47}{83} \approx 0.566 \text{ and } \frac{36}{83} \approx 0.434,$$

so these fractions give the same results we obtained with our sequence of vectors \mathbf{x}_k . These vectors provide a steady-state vector for Grand Rapids weather. In other words, if there is a 56.6% chance of it being dry on a given day in Grand Rapids, then there is a 56.6% chance it will be dry again the next day.

This is an example of a Markov process. Markov processes (named after Andrei Andreevich Markov) are widely used to model phenomena in biology, chemistry, business, physics, engineering, the social sciences, and much more. More specifically,



Definition 9.4. A **Markov process** is a process in which the probability of the system being in a given state depends only on the previous state.

If \mathbf{x}_0 is a vector which represents the initial state of a Markov process, then there is a matrix T (the *transition matrix*) such that the state of the system after one iteration is given by the vector $T\mathbf{x}_0$. This produces a chain of state vectors $T\mathbf{x}_0$, $T^2\mathbf{x}_0$, $T^3\mathbf{x}_0$, etc., where the state of the system after n iterations is given by $T^n\mathbf{x}_0$. Such a chain of vectors is called a *Markov chain*. A Markov process is characterized by two properties:

- the total number of observations remains fixed (this is reflected in the fact that the sum of the entries in each column of the matrix T is 1), and
- no observation is lost (this means the entries in the matrix T cannot be negative).

Summary

We learned about eigenvalues and eigenvectors of a matrix in this section.

- A scalar λ is an eigenvalue (or characteristic value) of a square matrix A if there is a non-zero vector \mathbf{x} so that $A\mathbf{x} = \lambda\mathbf{x}$.
- A non-zero vector \mathbf{x} is an eigenvector (or characteristic vector) of a square matrix A if there is a scalar λ so that $A\mathbf{x} = \lambda\mathbf{x}$.
- To find the eigenvectors of an $n \times n$ matrix A corresponding to an eigenvalue λ , we determine the non-trivial solutions to $(A - \lambda I_n)\mathbf{x} = \mathbf{0}$ where I_n is the $n \times n$ identity matrix.
- We study eigenvectors and eigenvalues because the eigenvectors tell us quite a bit about the transformation corresponding to the matrix. These eigenvectors arise in many applications in physics, chemistry, statistics, economics, biology, sociology and other areas, and help understand the long-term behavior of dynamical systems.
- A dynamical system is a system of variables whose values change with time. In linear dynamical systems, the change in the state vector from one time period to the next is expressed by matrix multiplication by the transition matrix A . The eigenvectors of A provide us a simple method to express the state vector at any given time period in terms of the initial state vector. Specifically, if the initial state vector is $\mathbf{x}_0 = a\mathbf{v}_0 + b\mathbf{w}_0$ where $\mathbf{v}_0, \mathbf{w}_0$ are eigenvectors corresponding to eigenvalues λ_1, λ_2 , we have

$$\mathbf{x}_k = A^k \mathbf{x}_0 = \lambda_1^k a \mathbf{v}_0 + \lambda_2^k b \mathbf{w}_0 .$$

Exercises

- (1) For each of the following matrix-vector pairs, determine whether the given vector is an eigenvector of the matrix.

(a) $A = \begin{bmatrix} 1 & 2 \\ 4 & 3 \end{bmatrix}, \mathbf{v} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$



$$(b) A = \begin{bmatrix} 1 & 2 \\ 0 & 3 \end{bmatrix}, \mathbf{v} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

$$(c) A = \begin{bmatrix} 2 & 1 & 0 \\ 0 & 1 & 0 \\ -1 & 0 & 1 \end{bmatrix}, \mathbf{v} = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$$

- (2) For each of the following matrix-eigenvalue pairs, determine an eigenvector of A for the given eigenvalue.

$$(a) A = \begin{bmatrix} 1 & 2 \\ -1 & 4 \end{bmatrix}, \lambda = 3 \quad (b) A = \begin{bmatrix} 1 & 4 \\ 1 & 1 \end{bmatrix}, \lambda = 3$$

$$(c) A = \begin{bmatrix} -1 & 4 & 1 \\ 3 & 3 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \lambda = 5 \quad (d) A = \begin{bmatrix} 4 & 0 & 0 & 0 \\ 0 & 2 & 0 & 2 \\ 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 3 \end{bmatrix}, \lambda = 4$$

- (3) For each of the following matrix- λ pairs, determine whether the given λ will work as an eigenvalue. You do not need to find an eigenvector as long you can justify if λ is a valid eigenvalue or not.

$$(a) A = \begin{bmatrix} 4 & 3 \\ 4 & 8 \end{bmatrix}, \lambda = 2 \quad (b) A = \begin{bmatrix} 4 & -2 \\ 2 & -1 \end{bmatrix}, \lambda = 0$$

$$(c) A = \begin{bmatrix} 1 & 2 \\ -1 & 4 \end{bmatrix}, \lambda = -1 \quad (d) A = \begin{bmatrix} 0 & -2 \\ -1 & 1 \end{bmatrix}, \lambda = -2$$

- (4) For a matrix A with eigenvector $\mathbf{v}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ with eigenvalue $\lambda_1 = 2$, and eigenvector $\mathbf{v}_2 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ with eigenvalue $\lambda_2 = -1$, determine the value of the following expressions using matrix-vector product properties:

$$(a) A(2\mathbf{v}_1 + 3\mathbf{v}_2)$$

$$(b) A(A(\mathbf{v}_1 + 2\mathbf{v}_2))$$

$$(c) A^{20}(4\mathbf{v}_1 - 2\mathbf{v}_2)$$

- (5) In this problem we consider a discrete dynamical system that forms what is called a *Markov chain* (see Definition 9.4) which models the number of students attending and skipping a linear algebra class in a semester. Assume the course starts with 1,000,000 students on day 0. For any given class day, 90% of the students who attend a class attend the next class (and 10% of these students skip next class) while only 30% of those absent are there the next time (and 70% of these students continue skipping class).

- (a) We know that there will be 900,000 students in class on the second day and 100,000 students skipping class. On the third day, 90% of the 900,000 students (attenders) and 30% of the 100,000 students (skippers) will come back to class. Therefore, 840,000 students will attend class on the third day. On the other hand, 10% of 900,000 students and 70% of 100,000 students skip class on the third day, for a total of 160,000 students skipping class.

We can use variables to represent these numbers. Let a_n represent the number of students attending class n days after first day. So $a_0 = 1,000,000$, $a_1 = 900,000$, $a_2 = 840,000$. Let s_n represent the students skipping class. So $s_0 = 0$, $s_1 = 100,000$, $s_2 = 160,000$. Find a_3, s_3, a_4, s_4 .

- (b) Find a linear expression for a_{k+1} in terms of the previous day values, a_k and s_k , using the story given in the problem. Similarly, express s_{k+1} in terms of a_k and s_k .

- (c) Let \mathbf{x}_k represent the state vector: $\mathbf{x}_k = \begin{bmatrix} a_k \\ s_k \end{bmatrix}$. It describes the state of the whole system (students attending class and skipping class) in one vector. For example, $\mathbf{x}_0 = \begin{bmatrix} 1,000,000 \\ 0 \end{bmatrix}$ is the initial state. The state next day is $\mathbf{x}_1 = \begin{bmatrix} 900,000 \\ 100,000 \end{bmatrix}$.

Using your answer to the previous part, find a matrix A which describes how the system changes from one day to the other so that $\mathbf{x}_{k+1} = A\mathbf{x}_k$.

- (6) Label each of the following statements as True or False. Provide justification for your response.

- (a) **True/False** The number 0 cannot be an eigenvalue.
- (b) **True/False** The 0 vector cannot be an eigenvector.
- (c) **True/False** If \mathbf{v} is an eigenvector of A , then so is $2\mathbf{v}$.
- (d) **True/False** If \mathbf{v} is an eigenvector of A , then it is also an eigenvector of A^2 .
- (e) **True/False** If \mathbf{v} and \mathbf{u} are eigenvectors of A with the same eigenvalue, then $\mathbf{v} + \mathbf{u}$ is also an eigenvector with the same eigenvalue.
- (f) **True/False** If λ is an eigenvalue of A , then λ^2 is an eigenvalue of A^2 .
- (g) **True/False** A projection matrix satisfies $P^2 = P$. If P is a projection matrix, then the eigenvalues of P can only be 0 and 1.
- (h) **True/False** If λ is an eigenvalue of an $n \times n$ matrix A , then $1 + \lambda$ is an eigenvalue of $I_n + A$.
- (i) **True/False** If λ is an eigenvalue of two matrices A and B of the same size, then λ is an eigenvalue of $A + B$.
- (j) **True/False** If \mathbf{v} is an eigenvector of two matrices A and B of the same size, then it is also an eigenvector of $A + B$.
- (k) **True/False** A matrix A has 0 as an eigenvalue if and only if A has linearly dependent columns.

Project: Understanding the PageRank Algorithm

Sergey Brin and Lawrence Page, the founders of Google, decided that the importance of a web page can be judged by the number of links to it as well as the importance of those pages. It is this



idea that leads to the PageRank algorithm.² Google uses this algorithm (and others) to order search engine results. According to Google:³

PageRank works by counting the number and quality of links to a page to determine a rough estimate of how important the website is. The underlying assumption is that more important websites are likely to receive more links from other websites.

To rank how “important” a website is, we need to make some assumptions. We assume that a person visits a page and then surfs the web by selecting a link from that page – all links on a given page are assigned the same probability of being chosen. As an example, assume a small set of seven pages 1, 2, 3, 4, 5, 6, and 7 with links between the pages given by the arrows as shown in Figure 9.1.⁴ So, for example, there is a hyperlink from page 4 to page 3, but no hyperlink in the opposite direction. If a web surfer starts on page 5, then there is probability of $\frac{1}{2}$ that this person will surf to page 6 and a probability of $\frac{1}{2}$ that the surfer will move to page 4. If there is no link leaving a page, as in the case of page 3, then the probability of remaining there is 1.

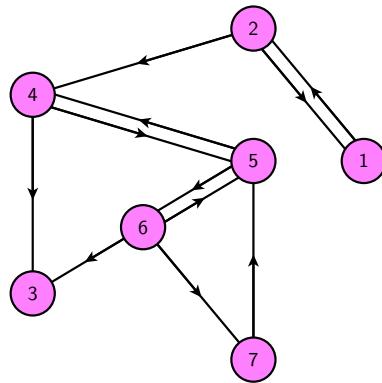


Figure 9.1: A seven page internet

To rank pages, we need to know how likely it is that a surfer will land on a given page. In our seven page example, a person can land on page 3 from page 4 with a probability of $\frac{1}{2}$ or from page 6 with a probability of $\frac{1}{3}$. If there is a link from a page we assume that the surfer leaves the page, and if there are no links from a page then the surfer stays on that page. We also assume that the surfer does not use the “Back” key. This information for our seven page internet example can be summarized in a *transition matrix* T whose i, j th entry is the probability that a surfer lands on page

²Information for this project was taken from the websites <http://www.ams.org/samplings/feature-column/fcarc-pagerank> and <http://faculty.winthrop.edu/polaskit/spring11/math550/chapter.pdf>.

³<http://web.archive.org/web/20111104131332/http://www.google.com/competition/howgooglesearchworks.html>

⁴The Internet is very large and has upwards of 25 billion pages. This would leave us with an enormous transition matrix, even though most of its entries are 0. In fact, studies show that web pages have an average of about 10 links, so on average all but 10 entries of each column are 0. Working with such a large matrix is beyond what we want to do in this project, so we will just amuse ourselves with small examples that illustrate the general points.

i from page j .

$$T = \begin{bmatrix} 0 & \frac{1}{2} & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & \frac{1}{2} & 0 & \frac{1}{3} & 0 \\ 0 & \frac{1}{2} & 0 & 0 & \frac{1}{2} & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{2} & 0 & \frac{1}{3} & 1 \\ 0 & 0 & 0 & 0 & \frac{1}{2} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \frac{1}{3} & 0 \end{bmatrix}.$$

Let us assume in our seven page internet that a user starts on page 6. That is, the probability that the user is initially on page 6 is 1, and so the probability that the user is on some other page is 0. This information can be encapsulated in a *state vector*

$$\mathbf{x}_0 = [0 \ 0 \ 0 \ 0 \ 1 \ 0]^T.$$

Since there are links from page 6 to pages 3, 5, and 7, there is a $\frac{1}{3}$ probability that the surfer will next move to one of these pages. That means that at the next step, the state vector \mathbf{x}_1 for this user will be

$$\mathbf{x}_1 = \left[0 \ 0 \ \frac{1}{3} \ 0 \ \frac{1}{3} \ 0 \ \frac{1}{3} \right]^T.$$

Note that

$$\mathbf{x}_1 = T\mathbf{x}_0.$$

As the user continues to surf the internet, the probabilities that the surfer is on a given page after the second, third, and fourth steps are given in the state vectors

$$\mathbf{x}_2 = T\mathbf{x}_1 = T^2\mathbf{x}_0, \quad \mathbf{x}_3 = T\mathbf{x}_2 = T^3\mathbf{x}_0, \quad \mathbf{x}_4 = T\mathbf{x}_3 = T^4\mathbf{x}_0.$$

In general, the probabilities that the surfer is on a given page after the n th step is given by the state vector

$$\mathbf{x}_n = T\mathbf{x}_{n-1} = T^n\mathbf{x}_0.$$

This example illustrates the general nature of what is called a *Markov process* (see Definition 9.4). The two properties of the transition matrix T make T a special kind of matrix.

Definition 9.5. A **stochastic** matrix is a matrix in which entries are nonnegative and the sum of the entries in every column is one.

In a Markov process, each generation depends only on the preceding generation and there may be a limiting value as we let the process continue indefinitely. We can test to see if that happens for this Markov process defined by T by doing some experimentation.

Project Activity 9.1. Use appropriate technology to do the following. Choose several different initial state vectors \mathbf{x}_0 and calculate the vectors in the sequence $\{T^n\mathbf{x}_0\}$ for large values of n . (Note that, as state vectors, the entries of \mathbf{x}_0 cannot be negative and the sum of the entries of \mathbf{x}_0 must be 1.) Explain the behavior of the sequence $\{\mathbf{x}_n\}$ as n gets large. Do you notice anything strange? What aspects of our seven page internet do you think explain this behavior? Clearly communicate all of the experimentation that you do. You may use the GeoGebra applet at <https://www.geogebra.org/m/b3dybnux>.



If there is a limit of the sequence $\{T^n \mathbf{x}_0\}$ (in other words, if there is a vector \mathbf{v} such that $\mathbf{v} = \lim_{n \rightarrow \infty} T^n \mathbf{x}_0$), we call this limit a *steady-state* or *equilibrium* vector. Such a steady-state vector has another important property. Since T is independent of n we have

$$T\mathbf{v} = T \left(\lim_{n \rightarrow \infty} T^n \mathbf{x}_0 \right) = \lim_{n \rightarrow \infty} T^{n+1} \mathbf{x}_0 = \mathbf{v}. \quad (9.8)$$

Equation (9.8) shows that a steady state vector \mathbf{v} is an eigenvector for T with eigenvalue 1. We can interpret the steady-state vector for T in an important way. Let t_j be the fraction of time we spend on page j and let l_j be the number of links on page j . Then the fraction of the time that we end up on page i coming from page j is $\frac{t_j}{l_j}$. If we sum over all the pages linked to page i we have that

$$t_i = \sum \frac{t_j}{l_j}.$$

Notice that this is essentially the same process we used to obtain \mathbf{x}_n from \mathbf{x}_{n-1} , and so we can interpret the steady-state vector \mathbf{v} as telling us what fraction of a random web surfer's time was spent at each web page. If we assume that the time spent at a web page is a measure of its importance, then the steady-state vector tells us the relative importance of each web page. So this steady-state vector provides the page rankings for us. In other words,

The importance of a webpage may be measured by the relative size of the corresponding entry in the steady-state vector for an appropriately chosen Markov chain.

Project Activity 9.2. Show that the limiting vector you found in Project Activity 9.1 is an eigenvector of T with eigenvalue 1.

Project Activity 9.1 illustrates one problem with our seven page internet. The steady-state vector shows that page 3 is the only important page, but that hardly seems reasonable in the example since there are other pages that must have some importance. The problem is that page 3 is a “dangling” page and does not lead anywhere. Once a surfer reaches that page, they are stuck there, overemphasizing its importance. So this dangling page acts like a sink, ultimately drawing all surfers to it. To adjust for dangling pages, we make the assumption that if a surfer reaches a dangling page (one with no links emanating from it), the surfer will jump to any page on the web with equal probability. So in our seven page example, once a surfer reaches page 3 the surfer will jump to any page on the web with probability $\frac{1}{7}$.

Project Activity 9.3.

- (a) Determine the transition matrix for our seven page internet with this adjustment.
- (b) Approximate the steady-state vector for this adjusted matrix so that the entries are accurate to four decimal places. Use any appropriate technology to row reduce matrices.
- (c) According to this adjusted model, which web page is now the most important? Why? Does this seem reasonable? Why?

There is one more issue to address before we can consider ourselves ready to rank web pages. Consider the example of the five page internet shown in Figure 9.2.



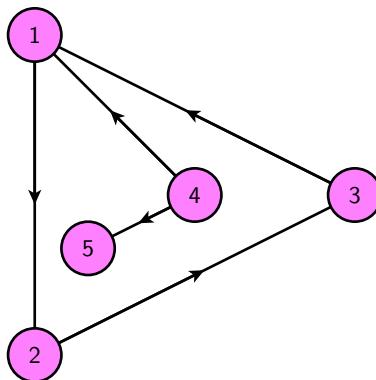


Figure 9.2: A five page internet

Project Activity 9.4.

- (a) Explain why

$$\begin{bmatrix} 0 & 0 & 1 & \frac{1}{2} & \frac{1}{5} \\ 1 & 0 & 0 & 0 & \frac{1}{5} \\ 0 & 1 & 0 & 0 & \frac{1}{5} \\ 0 & 0 & 0 & 0 & \frac{1}{5} \\ 0 & 0 & 0 & \frac{1}{2} & \frac{1}{5} \end{bmatrix}.$$

is the transition matrix for this five page internet. (Keep in mind the adjustment we made for dangling pages.)

- (b) Start with different initial state vectors
- \mathbf{x}_0
- and determine if there is a limit to the Markov chain. Explain. You may use the GeoGebra applet at
- <https://www.geogebra.org/m/b3dybnux>
- .

Project Activity 9.4 shows that it is possible to construct an internet so that the corresponding Markov chain does not have a limit, even after adjusting for dangling pages. This is a significant problem if we want to provide a relative ranking of all web pages regardless of where a surfer starts. To fix this problem we need to make one final adjustment to arrive at a type of transition matrix that always provides a limit for our Markov chain.

Definition 9.6. A stochastic matrix is **regular** if its transition matrix T has the property that for some power k , all the entries of T^k are positive.

Note that the transition matrix from Project Activity 9.4 is not regular. Regular matrices have some especially nice properties, as the following theorem describes. We will not prove this theorem, but use it in the remainder of this project. The theorem shows that if we have a regular transition matrix, then there will a limit of the state vectors \mathbf{x}_n , and that this limit has a very interesting property.

Theorem 9.7. Assume $n \geq 2$ and that T is a regular $n \times n$ stochastic matrix.

- (1) $\lim_{k \rightarrow \infty} T^k$ exists and is a stochastic matrix.

(2) For any vector \mathbf{x} ,

$$\lim_{k \rightarrow \infty} T^k \mathbf{x} = \mathbf{c}$$

for the same vector \mathbf{c} .

(3) The columns of $\lim_{k \rightarrow \infty} T^k$ are the same vector \mathbf{c} .

(4) The vector \mathbf{c} is the unique eigenvector of T whose entries sum to 1.

(5) If λ is an eigenvalue of T not equal to 1, then $|\lambda| < 1$.

Having a regular transition matrix T ensures that there is always the same limit \mathbf{v} to the sequence $T^k \mathbf{x}_0$ for any starting vector \mathbf{x}_0 . As mentioned before, the entries in $\mathbf{v} = \lim_{n \rightarrow \infty} T^n \mathbf{x}_0$ can be interpreted as telling us what fraction of the random surfer's time was spent at each webpage. If we interpret the amount of time a surfer spends at a page as a measure of the page's importance, then this steady-state vector \mathbf{v} provides a ranking of the relative importance of each page in the web. This is the essence of Google's PageRank.

To make our final adjustment in the transition matrix to be sure that we obtain a regular matrix, we need to deal with the problems of “loops” in our internet. Loops, as illustrated in Project Activity 9.4, can act as sinks just like the dangling pages we saw earlier and condemn a user that enters such a loop to spend his/her time only on those pages in the loop. Quite boring! To account for this problem, we make a second adjustment.

Let p be a number between 0 and 1 (Google supposedly uses $p = 0.85$). Suppose a surfer is on page i . We assume with probability p that the surfer will choose any link on page i with equal probability. We make the additional assumption with probability $1 - p$ that the surfer will select with equal probability any page on the web.

If T is a transition matrix, incorporating the method we used to deal with dangling pages, then the adjusted transition matrix G (the Google matrix) is

$$G = pT + (1 - p)Q,$$

where Q is the matrix all of whose entries are $\frac{1}{n}$, where n is the number of pages in the internet ($n = 7$ in our seven page example). Since all of the entries of G are positive, G is a regular stochastic matrix.

Project Activity 9.5. Return to the seven page internet in Figure 9.1.

- (a) Find the Google matrix G for this internet.
- (b) Approximate, to four decimal places, the steady-state vector for this internet.
- (c) What is the relative rank of each page in this internet, and approximately what percentage of time does a random user spend on each page?

We conclude with two observations. Consider the role of the parameter p in our final adjustment. Notice that if $p = 1$, then $G = T$ and we have the original hyperlink structure of the web. However, if $p = 0$, then $G = \frac{1}{n} I_n$, where I_n is the $n \times n$ identity matrix with n as the number of pages in the web. In this case, every page is linked to every other page and a random surfer spends equal time



on any page. Here we have lost all of the character of the linked structure of the web. Choosing p close to 1 retains much of the original hyperlink structure of the web.

Finally, the matrices that model the web are HUGE, and so the methods we used in this project to approximate the steady-state vectors are not practical. There are many methods for approximating eigenvectors that are often used in these situations, some of which we discuss in a later section.



Section 10

The Inverse of a Matrix

Focus Questions

By the end of this section, you should be able to give precise and thorough answers to the questions listed below. You may want to keep these questions in mind to focus your thoughts as you complete the section.

- What does it mean for a matrix A to be invertible?
- How can we tell when an $n \times n$ matrix A is invertible?
- If an $n \times n$ matrix A is invertible, how do we find the inverse of A ?
- If A and B are invertible $n \times n$ matrices, why is AB invertible and what is $(AB)^{-1}$?
- How can we use the inverse of a matrix in solving matrix equations?

Application: Modeling an Arms Race

Lewis Fry Richardson was a Quaker by conviction who was deeply troubled by the major wars that had been fought in his lifetime. Richardson's training as a physicist led him to believe that the causes of war were phenomena that could be quantified, studied, explained, and thus controlled. He collected considerable data on wars and constructed a model to represent an arms race. The equations in his model caused him concern about the future as indicated by the following statement:

But it worried him that the equations also showed that the unilateral disarmament of Germany after 1918, enforced by the Allied Powers, combined with the persistent level of armaments of the victor countries would lead to the level of Germanys armaments growing again. In other words, the post-1918 situation was not stable. From the model he concluded that great statesmanship would be needed to prevent an unstable situation from developing, which could only be prevented by a change of policies.¹

¹Nature 135, 830-831 (18 May 1935) "Mathematical Psychology of War" (3420).

Analyzing Richardson's arms race model utilizes matrix operations, including matrix inverses. We explore the basic ideas in Richardson's model later in this section.

Introduction

To this point we have solved systems of linear equations with matrix forms $Ax = \mathbf{b}$ by row reducing the augmented matrices $[A \mid \mathbf{b}]$. These linear matrix-vector equations should remind us of linear algebraic equations of the form $ax = b$, where a and b are real numbers. Recall that we solved an equation of the form $ax = b$ by dividing both sides by a (provided $a \neq 0$), giving the solution $x = \frac{b}{a}$, or equivalently $x = a^{-1}b$. The important property that the number a^{-1} has that allows us to solve a linear equation in this way is that $a^{-1}a = 1$, so that a^{-1} is the multiplicative inverse of a . We can solve certain types of matrix equations $Ax = \mathbf{b}$ in the same way, provided we can find a matrix A^{-1} with similar properties. We investigate this situation in this section.

Preview Activity 10.1.

- (1) Before we define the inverse matrix, recall that the identity matrix I_n (with 1's along the diagonal and 0's everywhere else) is a multiplicative identity in the set of $n \times n$ matrices (just like the real number 1 is the multiplicative identity in the set of real number). In particular, $I_n A = A I_n = A$ for any $n \times n$ matrix A .

Now we can generalize the inverse operation to matrices. For an $n \times n$ matrix A , we define A^{-1} to be the matrix which when multiplied by A gives us the identity matrix. In other words, $AA^{-1} = A^{-1}A = I_n$. We can find the inverse of a matrix in a calculator by using the x^{-1} button.

For each of the following matrices, determine if the inverse exists using your calculator or other appropriate technology. If the inverse does exist, write down the inverse and check that it satisfies the defining property of the inverse matrix, that is $AA^{-1} = A^{-1}A = I_n$. If the inverse doesn't exist, write down any error you received from the technology. Can you guess why the inverse does not exist for these matrices?

$$\begin{array}{ll} \text{(a)} \quad A = \begin{bmatrix} 1 & 3 \\ 0 & 4 \end{bmatrix} & \text{(b)} \quad A = \begin{bmatrix} 2 & 3 \\ 4 & 6 \end{bmatrix} \\ \text{(c)} \quad A = \begin{bmatrix} 1 & 2 & 3 \\ -1 & -1 & 2 \\ 1 & 2 & 2 \end{bmatrix} & \text{(d)} \quad A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 6 \\ 1 & 2 & 2 \end{bmatrix} \\ \text{(e)} \quad A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix} & \text{(f)} \quad A = \begin{bmatrix} 1 & 2 & 3 \\ -1 & -1 & 2 \\ 0 & 1 & 5 \end{bmatrix} \end{array}$$

- (2) Now we turn to the question of how to find the inverse of a matrix in general. With this approach, we will be able to determine which matrices have inverses as well.

We will consider the 2×2 case to make the calculations easier. Suppose A is a 2×2 matrix. Our goal is to find a matrix B so that $AB = I_2$ and $BA = I_2$. If such a matrix exists, we will call B the inverse, A^{-1} , of A .

- (a) What does the equation $AB = I_2$ tell us about the size of the matrix B ?



- (b) Now let $A = \begin{bmatrix} 1 & 2 \\ 1 & 3 \end{bmatrix}$. We want to find a matrix B so that $AB = I_2$. Suppose B has columns \mathbf{b}_1 and \mathbf{b}_2 , i.e. $B = [\mathbf{b}_1 \ \mathbf{b}_2]$. Our definition of matrix multiplication shows that

$$AB = [A\mathbf{b}_1 \ A\mathbf{b}_2].$$

- i. If $AB = I_2$, what must $A\mathbf{b}_1$ and $A\mathbf{b}_2$ equal?
 - ii. Use the result from part (a) to set up two matrix equations to solve to find \mathbf{b}_1 and \mathbf{b}_2 . Then find \mathbf{b}_1 and \mathbf{b}_2 . As a result, find the matrix B .
 - iii. When we solve the two systems we have found a matrix B so that $AB = I_2$. Is this enough to say that B is the inverse of A ? If not, what else do we need to know to verify that B is in fact A^{-1} ? Verify that B is A^{-1} .
- (3) A matrix inverse is extremely useful in solving matrix equations and can help us in solving systems of equations. Suppose that A is an invertible matrix, i.e., there exists A^{-1} such that $AA^{-1} = A^{-1}A = I_n$.
- (a) Consider the system $A\mathbf{x} = \mathbf{b}$. Use the inverse of A to show that this system has a solution for every \mathbf{b} and find an expression for this solution in terms of \mathbf{b} and A^{-1} . (Note that since matrix multiplication is not commutative, we have to pay attention to the order in which we multiply matrices. For example, $A^{-1}AB = B$ while we cannot simplify ABA^{-1} to B unless A and B commute.)
 - (b) If A , B , and C are matrices and $A + C = B + C$, then we can subtract the matrix C from both sides to see that $A = B$. We saw in Section 8 that there is no corresponding general cancellation property for matrix multiplication when we found that $AB = AC$ could hold while $B \neq C$. However, we can cancel A from this equation in certain circumstances. Suppose that $AB = AC$ and that A is an invertible matrix. Show that we can cancel A in this case and conclude that $B = C$. (Note: When simplifying the product of matrices, again keep in mind that matrix multiplication is not commutative.)

Invertible Matrices

We now have an algebra of matrices in that we can add, subtract, and multiply matrices of the correct sizes. But what about division? In our early mathematics education we learned about *multiplicative inverses* (or reciprocals) of real numbers. The multiplicative inverse of a number a is the real number which when multiplied by a produces 1, the multiplicative identity of real numbers. This inverse is denoted a^{-1} . For example, the multiplicative inverse of 2 is $2^{-1} = \frac{1}{2}$ because

$$2 \cdot \frac{1}{2} = 1 = \frac{1}{2} \cdot 2.$$

Of course, we didn't have to write both products because multiplication of real numbers is a commutative operation. There are a couple of important things to note about multiplicative inverses – we can use the inverses of the number a to solve the simple linear equation $ax + b = c$ for x ($x = a^{-1}(c - b)$), and not every real number has an inverse. The latter means that the inverse is not



defined on the entire set of real numbers. We can extend the idea of inverses to matrices, although we will see that there are many more matrices than just the zero matrix that do not have inverses.

To define matrix inverses² we make an analogy with the property of inverses in the real numbers: $x \cdot x^{-1} = 1 = x^{-1} \cdot x$.

Definition 10.1. Let A be an $n \times n$ matrix.

- (1) A is **invertible** if there is an $n \times n$ matrix B so that $AB = BA = I_n$.
- (2) If A is invertible, an **inverse** of A is a matrix B such that $AB = BA = I_n$.

If an $n \times n$ matrix A is invertible, its inverse will be unique (see Exercise 1), and we denote the inverse of A as A^{-1} . We also call an invertible matrix a *non-singular* matrix (with *singular* meaning non-invertible).

Activity 10.1.

- (a) Let $A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$. Calculate AB where $B = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$. Using your result, explain why it is not possible to have $AB = I_2$, showing that A is non-invertible.
- (b) Calculate AB where $A = \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}$ and $B = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$. Using your result, explain why the inverse of A doesn't exist.

We saw in Activity 10.1 why the inverse does not exist for two specific matrices. We will find in the next section an easy criterion for determining when a matrix has an inverse. In short, when the RREF of the matrix has a pivot in every column and row, then the matrix will be invertible. We know that this condition relates to quite a few other linear algebra concepts we have seen so far, such as linear independence of columns and the columns spanning \mathbb{R}^n . We will put these criteria together in one big theorem in the next section.

Activity 10.2. Suppose that A is an invertible $n \times n$ matrix. Hence we have an inverse matrix A^{-1} for which $AA^{-1} = A^{-1}A = I_n$. We will see how the inverse is useful in solving matrix equations involving A .

- (a) Explain why the matrix expressions

$$A^{-1}(AB), A^{-1}(A(BA)A^{-1}) \text{ and } BA^{-1}BAA^{-1}B^{-1}A$$

can all be simplified to B . (Hint: Use the associative property of matrix multiplication.)

- (b) Suppose the system $Ax = b$ has a solution. Explain why then $A^{-1}(Ax) = A^{-1}b$. What does this equation simplify to?
- (c) Since we found one single expression for the solution x in equation $Ax = b$, this implies that the equation has a unique solution. What does this imply about the matrix A ?

²We usually refer to a multiplicative inverse as just an inverse. Since every matrix has an additive inverse, there is no need to consider the existence of additive inverses.



As we saw in Preview Activity 10.1, if the $n \times n$ matrix A is invertible, then the equation $A\mathbf{x} = \mathbf{b}$ is consistent for all \mathbf{b} in \mathbb{R}^n and has the unique solution $\mathbf{x} = A^{-1}\mathbf{b}$. This means that A has a pivot in every row and column, which is equivalent to the criterion that A reduces to I_n , as we noted above.

Even though $\mathbf{x} = A^{-1}\mathbf{b}$ is an explicit expression for the solution of the system $A\mathbf{x} = \mathbf{b}$, using the inverse of a matrix is usually not a computationally efficient way to solve a matrix equation. Finding the RREF of a matrix computationally takes fewer steps to solve the matrix equation.

Finding the Inverse of a Matrix

The next questions for us to address are how to tell when a matrix is invertible and how to find the inverse of an invertible matrix. Consider a 2×2 matrix A . To find the inverse matrix $B = [\mathbf{b}_1 \ \mathbf{b}_2]$ of A , we have to solve the two matrix-vector equations $A\mathbf{b}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $A\mathbf{b}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ to find the columns of B . Since A is the coefficient matrix for both systems, we apply the same row operations on both systems to reduce A to RREF. Thus, instead of solving the two matrix-vector equations separately, we could simply have found the RREF of

$$\left[A \mid \begin{array}{cc} 1 & 0 \\ 0 & 1 \end{array} \right]$$

and done all of the work in one pass. Note that the right hand side of the augmented matrix is now I_2 . So we row reduce $[A \mid I_2]$, and if the systems are consistent, the reduced row echelon form of $[A \mid I_2]$ must be $[I_2 \mid A^{-1}]$. You should be able to see that this same process works in any dimension.

How to find the inverse of an $n \times n$ matrix A :

- Augment A with the identity matrix I_n .
- Apply row operations to reduce the augmented matrix $[A \mid I_n]$. If the system is consistent, then the reduced row echelon form of $[A \mid I_n]$ will have the form $[I_n \mid B]$ (by Activity 10.1 (d)). If the reduced row echelon form of A is not I_n , then this step fails and A is not invertible.
- If A is row equivalent to I_n , then the matrix B in the second step has the property that $AB = I_n$. We will show later that the matrix B also satisfies $BA = I_n$ and so B is the inverse of A .

Activity 10.3. Find the inverse of each matrix using the method above, if it exists. Compare the result with the inverse that you get from using appropriate technology to directly calculate the inverse.

(a) $\left[\begin{array}{ccc} 1 & 1 & 1 \\ 1 & 1 & -1 \\ 1 & -1 & 0 \end{array} \right]$

(b) $\left[\begin{array}{ccc} 1 & 1 & 1 \\ 2 & 2 & 2 \\ 0 & 0 & 1 \end{array} \right]$

We can use this method of finding the inverse of a matrix to derive a concrete formula for the inverse of a 2×2 matrix:

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}, \quad (10.1)$$

provided that $ad - bc \neq 0$ (see Exercise 2). Hence, any 2×2 matrix $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ has an inverse if and only if $ad - bc \neq 0$. We call this quantity *determinant of A*, $\det(A)$. We will see that the determinant of a general $n \times n$ matrix will be essential in determining invertibility of the matrix.

Properties of the Matrix Inverse

As we have done with every new operation, we ask what properties the inverse of a matrix has.

Activity 10.4. Consider the following questions about matrix inverses. If two $n \times n$ matrices A and B are invertible, is the product AB invertible? If so, what is the inverse of AB ? We answer these questions in this activity.

(a) Let

$$A = \begin{bmatrix} 1 & 2 \\ 1 & 3 \end{bmatrix} \text{ and } B = \begin{bmatrix} 2 & 3 \\ -1 & 2 \end{bmatrix}.$$

- i. Use formula (10.1) to find the inverses of A and B .
- ii. Find the matrix product AB . Is AB invertible? If so, use formula (10.1) to find the inverse of AB .
- iii. Calculate the products $A^{-1}B^{-1}$ and $B^{-1}A^{-1}$. What do you notice?

(b) In part (a) we saw that the matrix product $B^{-1}A^{-1}$ was the inverse of the matrix product AB . Now we address the question of whether this is true in general. Suppose now that C and D are invertible $n \times n$ matrices so that the matrix inverses C^{-1} and D^{-1} exist.

- i. Use matrix algebra to simplify the matrix product $(CD)(D^{-1}C^{-1})$. (Hint: What do you know about DD^{-1} and CC^{-1} ?)
- ii. Simplify the matrix product $(D^{-1}C^{-1})(CD)$ in a manner similar to part i.
- iii. What conclusion can we draw from parts i and ii? Explain. What property of matrix multiplication requires us to reverse the order of the product when we create the inverse of CD ?

Activity 10.4 gives us one important property of matrix inverses. The other properties given in the next theorem can be verified similarly.

Theorem 10.2. *Let A and B be invertible $n \times n$ matrices. Then*

$$(1) \quad (A^{-1})^{-1} = A.$$

$$(2) \quad \text{The product } AB \text{ is invertible and } (AB)^{-1} = B^{-1}A^{-1}.$$

$$(3) \quad \text{The matrix } A^T \text{ is invertible and } (A^T)^{-1} = (A^{-1})^T.$$



Examples

What follows are worked examples that use the concepts from this section.

Example 10.3. For each of the following matrices A ,

- Use appropriate technology to find the reduced row echelon form of $[A | I_3]$.
- Based on the result of part (a), is A invertible? If yes, what is A^{-1} ? If no, explain why.
- Let $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$ and $\mathbf{b} = \begin{bmatrix} 5 \\ 4 \\ 1 \end{bmatrix}$. If A is invertible, solve the matrix equation $A\mathbf{x} = \mathbf{b}$ using the inverse of A . If A is not invertible, find all solutions, if any, to the equation $A\mathbf{x} = \mathbf{b}$ using whatever method you choose.

$$(a) A = \begin{bmatrix} 1 & 2 & 3 \\ 1 & -1 & -1 \\ 1 & 0 & 1 \end{bmatrix}$$

$$(b) A = \begin{bmatrix} 1 & 2 & 5 \\ 1 & -1 & -1 \\ 1 & 0 & 1 \end{bmatrix}$$

Example Solution.

$$(a) \text{ With } A = \begin{bmatrix} 1 & 2 & 3 \\ 1 & -1 & -1 \\ 1 & 0 & 1 \end{bmatrix}, \text{ we have the following.}$$

- The reduced row echelon form of $[A | I_3]$ is

$$\left[\begin{array}{ccc|ccc} 1 & 0 & 0 & \frac{1}{2} & 1 & -\frac{1}{2} \\ 0 & 1 & 0 & 1 & 1 & -2 \\ 0 & 0 & 1 & -\frac{1}{2} & -1 & \frac{3}{2} \end{array} \right].$$

- Since A is row equivalent to I_3 , we conclude that A is invertible. The reduced row echelon form of $[A | I_3]$ tells us that

$$A^{-1} = \frac{1}{2} \begin{bmatrix} 1 & 2 & -1 \\ 2 & 2 & -4 \\ -1 & -2 & 3 \end{bmatrix}.$$

- The solution to $A\mathbf{x} = \mathbf{b}$ is given by

$$v\mathbf{x} = A^{-1}\mathbf{b} = \frac{1}{2} \begin{bmatrix} 1 & 2 & -1 \\ 2 & 2 & -4 \\ -1 & -2 & 3 \end{bmatrix} \begin{bmatrix} 5 \\ 4 \\ 1 \end{bmatrix} = \begin{bmatrix} 6 \\ 7 \\ -5 \end{bmatrix}.$$



(b) With $A = \begin{bmatrix} 1 & 2 & 5 \\ 1 & -1 & -1 \\ 1 & 0 & 1 \end{bmatrix}$, we have the following.

- The reduced row echelon form of $[A | I_3]$ is

$$\left[\begin{array}{ccc|ccc} 1 & 0 & 1 & 0 & 0 & 1 \\ 0 & 1 & 2 & 0 & -1 & 1 \\ 0 & 0 & 0 & 1 & 2 & -3 \end{array} \right].$$

- Since A is not row equivalent to I_3 , we conclude that A is not invertible.
- The reduced row echelon form of $[A | \mathbf{b}]$ is

$$\left[\begin{array}{ccc|c} 1 & 0 & 1 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 1 \end{array} \right].$$

The fact that the augmented column is a pivot column means that the equation $A\mathbf{x} = \mathbf{b}$ has no solutions.

Example 10.4.

(a) Let $A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$.

- Show that $A^2 \neq 0$ but $A^3 = 0$.
 - Show that $I - A$ is invertible and find its inverse. Compare the inverse of $I - A$ to $I + A + A^2$.
- (b) Let M be an arbitrary square matrix such that $M^3 = 0$. Show that M is invertible and find an inverse for M .

Example Solution.

(a) Let $A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$.

- Using technology to calculate A^2 and A^3 we find that $A^3 = 0$ while $A^2 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$.

- For this matrix A we have $I - A = \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{bmatrix}$. The reduced row echelon form of $I - A$ is

$$\left[\begin{array}{ccc|ccc} 1 & -0 & 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 0 & 1 \end{array} \right],$$



so $I - A$ is invertible and $(I - A)^{-1} = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}$.

A straightforward matrix calculation also shows that

$$(I - A)^{-1} = I + A + A^2.$$

- (b) We can try to emulate the result of part (a) here. Expanding using matrix operations gives us

$$\begin{aligned} (I - M)(I + M + M^2) &= (I + M + M^2) - (M + M^2 + M^3) \\ &= (I + M + M^2) - (M + M^2 + 0) \\ &= I \end{aligned}$$

and

$$\begin{aligned} (I + M + M^2)(I - M) &= (I + M + M^2) - (M + M^2 + M^3) \\ &= (I + M + M^2) - (M + M^2 + 0) \\ &= I. \end{aligned}$$

So $I - M$ is invertible and $(I - M)^{-1} = I + M + M^2$.

This argument can be generalized to show that if M is a square matrix and $M^n = 0$ for some positive integer n , then $I - M$ is invertible and

$$(I - M)^{-1} = I + M + M^2 + \cdots + M^{n-1}.$$

Summary

- If A is an $n \times n$ matrix, then A is invertible if there is a matrix B so that $AB = BA = I_n$. The matrix B is called the inverse of A and is denoted A^{-1} .
- An $n \times n$ matrix A is invertible if and only if the reduced row echelon form of A is the $n \times n$ identity matrix I_n .
- To find the inverse of an invertible $n \times n$ matrix A , augment A with the identity and row reduce. If $[A | I_n] \sim [I_n | B]$, then $B = A^{-1}$.
- If A and B are invertible $n \times n$ matrices, then $(AB)^{-1} = B^{-1}A^{-1}$. Since the inverse of AB exists, the product of two invertible matrices is an invertible matrix.
- We can use the algebraic tools we have developed for matrix operations to solve equations much like we solve equations with real variables. We must be careful, though, to only multiply by inverses of invertible matrices, and remember that matrix multiplication is not commutative.



Exercises

- (1) Let A be an invertible $n \times n$ matrix. In this exercise we will prove that the inverse of A is unique. To do so, we assume that both B and C are inverses of A , that is $AB = BA = I_n$ and $AC = CA = I_n$. By considering the product BAC simplified in two different ways, show that $B = C$, implying that the inverse of A is unique.

- (2) Let $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ be an arbitrary 2×2 matrix.

- (a) If A is invertible, perform row operations to determine a row echelon form of A .
(Hint: You may need to consider different cases, e.g., when $a = 0$ and when $a \neq 0$.)
- (b) Under certain conditions, we can row reduce $[A | I_2]$ to $[I_2 | B]$, where

$$B = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}.$$

Use the row echelon form of A from part (a) to find conditions under which the 2×2 matrix A is invertible. Then derive the formula for the inverse B of A .

- (3) (a) For a few different k values, find the inverse of $A = \begin{bmatrix} 1 & k \\ 0 & 1 \end{bmatrix}$. From these results, make a conjecture as to what A^{-1} is in general.
- (b) Prove your conjecture using the definition of inverse matrix.

- (c) Find the inverse of $A = \begin{bmatrix} 1 & k & \ell \\ 0 & 1 & m \\ 0 & 0 & 1 \end{bmatrix}$.

(Note: You can combine the first two parts above by applying the inverse finding algorithm directly on $A = \begin{bmatrix} 1 & k \\ 0 & 1 \end{bmatrix}$.)

- (4) Solve for the matrix A in terms of the others in the following equation:

$$P^{-1}(D + CA)P = B$$

If you need to use an inverse, assume it exists.

- (5) For which c is the matrix $A = \begin{bmatrix} 1 & 2 & -1 \\ 2 & 1 & 1 \\ 1 & 5 & c \end{bmatrix}$ invertible?

- (6) For which c is the matrix $A = \begin{bmatrix} c & 2 \\ 3 & c \end{bmatrix}$ invertible?

- (7) Let A and B be invertible $n \times n$ matrices. Verify the remaining properties of Theorem 10.2.
That is, show that

(a) $(A^{-1})^{-1} = A$.

(b) The matrix A^T is invertible and $(A^T)^{-1} = (A^{-1})^T$.



- (8) Label each of the following statements as True or False. Provide justification for your response.
- (a) **True/False** If A is an invertible matrix, then for any two matrices B, C , $AB = AC$ implies $B = C$.
 - (b) **True/False** If A is invertible, then so is AB for any matrix B .
 - (c) **True/False** If A and B are invertible $n \times n$ matrices, then so is AB .
 - (d) **True/False** If A is an invertible $n \times n$ matrix, then the equation $Ax = \mathbf{b}$ is consistent for any \mathbf{b} in \mathbb{R}^n .
 - (e) **True/False** If A is an invertible $n \times n$ matrix, then the equation $Ax = \mathbf{b}$ has a unique solution when it is consistent.
 - (f) **True/False** If A is invertible, then so is A^2 .
 - (g) **True/False** If A is invertible, then it reduces to the identity matrix.
 - (h) **True/False** If a matrix is invertible, then so is its transpose.
 - (i) **True/False** If A and B are invertible $n \times n$ matrices, then $A + B$ is invertible.
 - (j) **True/False** If $A^2 = 0$, then $I + A$ is invertible.

Project: The Richardson Arms Race Model

How and why a nation arms itself for defense depends on many factors. Among these factors are the offensive military capabilities a nation deems its enemies have, the resources available for creating military forces and equipment, and many others. To begin to analyze such a situation, we will need some notation and background. In this section we will consider a two nation scenario, but the methods can be extended to any number of nations. In fact, after World War I, Richardson collected data and created a model for the countries Czechoslovakia, China, France, Germany, England, Italy, Japan, Poland, the USA, and the USSR.³

Let N_1 and N_2 represent 2 different nations. Each nation has some military capability (we will call this the *armament* of the nation) at time n (think of n as representing the year). Let $a_1(n)$ represent the armament of nation N_1 at time n , and $a_2(n)$ the armament of nation N_2 at time n . We could measure $a_i(n)$ in weaponry or dollars or whatever units make sense for armaments. The Richardson arms race model provides connections between the armaments of the two nations.

Project Activity 10.1. We continue to analyze a two nation scenario. Let us suppose that our two nations are Iran (nation N_1) and Iraq (nation N_2). In 1980, Iraq invaded Iran resulting in a long and brutal 8 year war. Richardson was interested in analyzing data to see if such wars could be predicted by the changes in armaments of each nation. We construct the two nation model in this activity.

During each time period every nation adds or subtracts from its armaments. In our model, we will consider three main effects on the changes in armaments: the defense effect, fatigue effect and

³The Union of Soviet Socialist Republics (USSR), headed by Russia, was a confederation of socialist republics in Eurasia. The USSR disbanded in 1991. Czechoslovakia was a sovereign state in central Europe that peacefully split into the Czech Republic and Slovakia in 1993.



the grievance effect. In this activity we will discuss each effect in turn and then create a model to represent a two nation arms race.

- We first consider the defense effect. In a two nation scenario, each nation may react to the potential threat implied by an arms buildup of the other nation. For example, if nation N_1 feels threatened by nation N_2 (think of South and North Korea, or Ukraine and Russia, for example), then nation N_2 's level of armament might cause nation N_1 to increase its armament in response. We will let δ_{12} represent this effect of nation N_2 's armament on the armament of nation N_1 . Nation N_1 will then increase (or decrease) its armament in time period n by the amount $\delta_{12}a_2(n - 1)$ based on the armament of nation N_2 in time period $n - 1$. We will call δ_{12} a *defense coefficient*.⁴
- Next we discuss the fatigue effect. Keeping a strong defense is an expensive and taxing enterprise, often exacting a heavy toll on the resources of a nation. For example, consider the fatigue that the U.S. experienced fighting wars in Iraq and Afghanistan, losing much hardware and manpower in these conflicts. Let δ_{ii} represent this *fatigue factor* on nation i . Think of δ_{ii} as a measure of how much the nation has to replace each year, so a positive fatigue factor means that the nation is adding to its armament. The fatigue factor produces an effect of $\delta_{ii}a_i(n - 1)$ on the armament of nation i at time $t = n$ that is the effect of the armament at time $t = n - 1$.
- The last factor we consider is what we will call a grievance factor. This can be thought of as the set of ambitions and/or grievances against other nations (such as the acquisition or reacquisition of territory currently belonging to another country). As an example, Argentina and Great Britain both claim the Falkland Islands as territory. In 1982 Argentina invaded the disputed Falkland Islands which resulted in a two-month long undeclared Falkland Islands war, which returned control to the British. It seems reasonable that one nation might want to have sufficient armament in place to support its claim if force becomes necessary. Assuming that these grievances and ambitions have a constant impact on the armament of a nation from year to year, let g_i be this “grievance” constant for nation i .⁵ The effect a grievance factor g_i would have on the armament of nation i in year n would be to add g_i directly to $a_i(n - 1)$, since the factor g_i is constant from year to year (paying for arms and soldier's wages, for example) and does not depend on the amount of existing armament.

- (a) Taking the three effects discussed above into consideration, explain why

$$a_1(n) = \delta_{11}a_1(n - 1) + \delta_{12}a_2(n - 1) + a_1(n - 1) + g_1.$$

Then explain why

$$a_1(n) = (\delta_{11} + 1)a_1(n - 1) + \delta_{12}a_2(n - 1) + g_1. \quad (10.2)$$

⁴Of course, there are many other factors that have not been taken into account in the analysis. A nation may have heavily armed allies (like the U.S.) which may provide enough perceived security that this analysis is not relevant. Also, a nation might be a neutral state, such as Switzerland, and this analysis might not apply to such nations.

⁵It might be possible for g_i to be negative if, for example, a nation feels that such disputes can and should only be settled by negotiation.



(b) Write an equation similar to equation (10.2) that describes $a_2(n)$ in terms of the three effects.

(c) Let $\mathbf{a}_n = \begin{bmatrix} a_1(n) \\ a_2(n) \end{bmatrix}$. Explain why

$$\mathbf{a}_n = (D + I_2)\mathbf{a}_{n-1} + \mathbf{g},$$

where $D = \begin{bmatrix} \delta_{11} & \delta_{12} \\ \delta_{21} & \delta_{22} \end{bmatrix}$ and $\mathbf{g} = [g_1 \ g_2]^\top$.

Year	Iran	Iraq
1966	662	391
1967	903	378
1968	1090	495
1969	1320	615
1970	1470	600
1971	1970	618
1972	2500	589
1973	2970	785
1974	5970	2990
1975	7100	1690

Table 10.1: Military Expenditures of Iran and Iraq 1966-1975.

Project Activity 10.2. In order to analyze a specific arms race between nations, we need some data to determine values of the δ_{ij} and the g_i . Table 10.1 shows the military expenditures of Iran and Iraq in the years leading up to their war in 1975. (The data is in millions of US dollars, adjusted for inflation and is taken from “World Military Expenditures and Arms Transfers 1966-1975” by the U.S. Arms Control and Disarmament Agency.) We can perform regression (we will see how in a later section) on this data to obtain the following linear approximations:

$$a_1(n) = 2.0780a_1(n-1) - 1.7081a_2(n-1) - 126.9954 \quad (10.3)$$

$$a_2(n) = 0.9419a_1(n-1) - 1.3283a_2(n-1) - 101.2980. \quad (10.4)$$

(Of course, the data does not restrict itself to only factors between the two countries, so our model will not be as precise as we might like. However, it is a reasonable place to start.) Use the regression equations (10.3) and (10.4) to explain why

$$D = \begin{bmatrix} 1.0780 & -1.7081 \\ 0.94194 & -2.3283 \end{bmatrix} \text{ and } \mathbf{g} = [-126.9954 \ -101.2980]^\top$$

for our Iran-Iraq arms race.

Activities 10.1 and 10.2 provide the basics to describe the general arms race model due to Richardson. If we have an m nation arms race with $D = [\delta_{ij}]$ and $\mathbf{g} = [g_i]$, then

$$\mathbf{a}_n = (D + I_m)\mathbf{a}_{n-1} + \mathbf{g}. \quad (10.5)$$



Project Activity 10.3. The idea of an arms race, theoretically, is to reach a point at which all parties feel secure and no additional money needs to be spent on armament. If such a situation ever arises, then the armament of all nations is stable, or in equilibrium. If we have an equilibrium solution, then for large values of n we will have $\mathbf{a}_n = \mathbf{a}_{n-1}$. So to find an equilibrium solution, if it exists, we need to find a vector \mathbf{a}_E so that

$$\mathbf{a}_E = (D + I)\mathbf{a}_E + \mathbf{g} \quad (10.6)$$

where I is the appropriate size identity matrix. If \mathbf{a}_E exists, we call \mathbf{a}_E an *equilibrium state*.

We can apply matrix algebra to find the equilibrium state vector \mathbf{a}_E under certain conditions.

- (a) Assuming that \mathbf{a}_E exists, use matrix algebra and Equation 10.6 to show that

$$D\mathbf{a}_E + \mathbf{g} = 0. \quad (10.7)$$

- (b) Under what conditions can we be assured that there will always be a unique equilibrium state \mathbf{a}_E ? Explain. Under these conditions, how can we find this unique equilibrium state? Write this equilibrium state vector \mathbf{a}_E as a matrix-vector product.
- (c) Does the arms race model for Iran and Iraq have an equilibrium solution? If so, find it. If not, explain why not. Use technology as appropriate.
- (d) Assuming an equilibrium exists and that both nations behave in a way that supports the equilibrium, explain what the appropriate entry of the equilibrium state vector \mathbf{a}_E suggests about what Iran and Iraq's policies should be. What does this model say about why there might have been war between these two nations?

Section 11

The Invertible Matrix Theorem

Focus Questions

By the end of this section, you should be able to give precise and thorough answers to the questions listed below. You may want to keep these questions in mind to focus your thoughts as you complete the section.

- What does it mean for two statements to be equivalent?
- How can we efficiently prove that a string of statements are all equivalent?
- What is the Invertible Matrix Theorem and why is it important?
- What are the equivalent conditions to a matrix being invertible?

Introduction

This section is different than others in this book in that it contains only one long proof of the equivalence of statements that we have already discussed. As such, this is a theoretical section and there is no application connected to it.

The Invertible Matrix Theorem is a theorem that provides many different statements that are equivalent to having a matrix be invertible. To understand the Invertible Matrix Theorem, we need to know what it means for two statements to be *equivalent*. By equivalent, we mean that if one of the statements is true, then so is the other. We examine this idea in this preview activity.

Preview Activity 11.1. Let A be an $n \times n$ matrix. In this activity we endeavor to understand why the two statements

- I. The matrix A is invertible.
- II. The matrix A^T is invertible.

are equivalent. To demonstrate that statements I and II are equivalent, we need to argue that if statement I is true, then so is statement II, and if statement II is true then so is statement I.

- (1) Let's first show that if statement I is true, then so is statement II. So we assume statement I. That is, we assume that A is an invertible matrix. So we know that there is an $n \times n$ matrix B such that $AB = BA = I_n$, where I_n is the $n \times n$ identity matrix. To demonstrate that statement II must also be true, we need to verify that A^T is an invertible matrix.
- What is I_n^T ?
 - Take the transpose of both sides of the equation $AB = I_n$ and use the properties of the transpose to write $(AB)^T$ in terms of A^T and B^T .
 - Take the transpose of both sides of the equation $BA = I_n$ and use the properties of the transpose to write $(BA)^T$ in terms of A^T and B^T .
 - Explain how the previous two parts show that B^T is the inverse of A^T , so that A^T is invertible. So we have shown that if statement I is true, so is statement II.¹
- (2) Now we prove that if statement II is true, then so is statement I. So we assume statement II. That is, we assume that the matrix A^T is invertible. We could do this in the same manner as part (a), or we could be a bit clever. Let's try to be clever.
- What matrix is $(A^T)^T$?
 - Why can we use the result of part (a) with A^T in place of A to conclude that A is invertible? As a consequence, we have demonstrated that A is invertible if A^T is invertible. This concludes our argument that statements I and II are equivalent.

The Invertible Matrix Theorem

We have been introduced to many statements about existence and uniqueness of solutions to systems of linear equations, linear independence of columns of coefficient matrices, onto linear transformations, and many other items. In this section we will analyze these statements in light of how they are related to invertible matrices, with the main goal to understand the Invertible Matrix Theorem.

Recall that an $n \times n$ matrix A is invertible if there is an $n \times n$ matrix B such that $AB = BA = I_n$, where I_n is the $n \times n$ identity matrix. The Invertible Matrix Theorem is an important theorem in that it provides us with a wealth of statements that are all equivalent to the statement that an $n \times n$ matrix A is invertible, and connects many of the topics we have been discussing so far this semester into one big picture.

Theorem 11.1 (The Invertible Matrix Theorem). *Let A be an $n \times n$ matrix. The following statements are equivalent:*

- (1) *A is an invertible matrix.*
- (2) *The equation $Ax = \mathbf{0}$ has only the trivial solution.*
- (3) *A has n pivot columns.*
- (4) *The columns of A span \mathbb{R}^n .*

¹Note that statement I does not have to be true. We are only assuming that IF statement I is true, then statement II must also be true.



- (5) *A is row equivalent to the identity matrix I_n .*
- (6) *The columns of A are linearly independent.*
- (7) *The columns of A form a basis for \mathbb{R}^n .*
- (8) *The matrix transformation T from \mathbb{R}^n to \mathbb{R}^n defined by $T(\mathbf{x}) = A\mathbf{x}$ is one-to-one.*
- (9) *The matrix equation $A\mathbf{x} = \mathbf{b}$ has exactly one solution for each vector \mathbf{b} in \mathbb{R}^n .*
- (10) *The matrix transformation T from \mathbb{R}^n to \mathbb{R}^n defined by $T(\mathbf{x}) = A\mathbf{x}$ is onto.*
- (11) *There is an $n \times n$ matrix C so that $AC = I_n$.*
- (12) *There is an $n \times n$ matrix D so that $DA = I_n$.*
- (13) *The scalar 0 is not an eigenvalue of A.*
- (14) *A^T is invertible.*

The Invertible Matrix Theorem is a theorem that provides many different statements that are equivalent to a matrix being invertible. As discussed in Preview Activity 11.1, two statements are said to be **equivalent** if, whenever one of the statements is true, then the other is also true. So to demonstrate, say, statements I and II are equivalent, we need to argue that

- if statement I is true, then so is statement II, and
- if statement II is true then so is statement I.

The Invertible Matrix Theorem, however, provides a long list of statements that are equivalent. It would be inefficient to prove, one by one, that each pair of statements is equivalent. (There are $\binom{14}{2} = 91$ such pairs.) Fortunately, there is a shorter method that we can use.

Activity 11.1. In this activity, we will consider certain parts of the Invertible Matrix Theorem and show that one implies another in a specific order. For all parts in this activity, we assume A is an $n \times n$ matrix.

- (a) Consider the following implication:
 $(2) \implies (6)$:² If the equation $A\mathbf{x} = \mathbf{0}$ has only the trivial solution, then the columns of A are linearly independent. This shows that part 2 of the IMT implies part 6 of the IMT. Justify this implication as if it is a T/F problem.
- (b) Justify the following implication:
 $(6) \implies (9)$: If the columns of A are linearly independent, then the matrix equation $A\mathbf{x} = \mathbf{b}$ has exactly one solution for each vector \mathbf{b} in \mathbb{R}^n .
- (c) Justify the following implication:
 $(9) \implies (4)$: If the equation $A\mathbf{x} = \mathbf{b}$ has exactly one solution for every vector \mathbf{b} in \mathbb{R}^n , then the columns of A span \mathbb{R}^n .

²The symbol \implies is the implication symbol, so $(1) \implies (12)$ is read to mean that statement (1) of the theorem implies statement (12).

- (d) Justify the following implication:
 $(4) \implies (2)$: If the columns of A span \mathbb{R}^n , then the equation $Ax = \mathbf{0}$ has only the trivial solution.
- (e) Using the above implications you proved, explain why we can conclude the following implication must also be true:
 $(2) \implies (9)$: If the equation $Ax = \mathbf{0}$ has only the trivial solution, then the matrix equation $Ax = \mathbf{b}$ has exactly one solution for each vector \mathbf{b} in \mathbb{R}^n .
- (f) Using the above implications you proved, explain why any one of the implications (2), (6), (9), and (4) implies any of the others.

Using a similar ordering of circular implications as in Activity 11.1, we can prove the Invertible Matrix Theorem by showing that each statement in the list implies the next statement, and that the last statement implies the first.

Proof of the Invertible Matrix Theorem

Statement (1) implies Statement (2). This follows from work done in Activity 11.1.

Statement (2) implies Statement (3). This was done in Activity 11.1.

Statement (3) implies Statement (4). Suppose that every column of A is a pivot column. The fact that A is square means that every row of A contains a pivot, and hence the columns of A span \mathbb{R}^n .

Statement (4) implies Statement (5). Since the columns of A span \mathbb{R}^n , it must be the case that every row of A contains a pivot. This means that A must be row equivalent to I_n .

Statement (5) implies Statement (6). If A is row equivalent to I_n , there must be a pivot in every column, which means that the columns of A are linearly independent.

Statement (6) implies Statement (7). If the columns of A are linearly independent, then there is a pivot in every column. Since A is a square matrix, there is a pivot in every row as well. So the columns of A span \mathbb{R}^n . Since they are also linearly independent, the columns form a minimal spanning set, which is a basis of \mathbb{R}^n .

Statement (7) implies Statement (8). If the columns form a basis of \mathbb{R}^n , then the columns are linearly independent. This means that each column is a pivot column, which also implies $Ax = \mathbf{0}$ has a unique solution and that T is one-to-one.

Statement (8) implies Statement (9). If T is one-to-one, then A has a pivot in every column. Since A is square, every row of A contains a pivot. Therefore, the system $Ax = \mathbf{b}$ is consistent for every $\mathbf{b} \in \mathbb{R}^n$ and has a unique solution.

Statement (9) implies Statement (10). If $Ax = \mathbf{b}$ has a unique solution for every \mathbf{b} , then the transformation T is onto since $T(\mathbf{x}) = \mathbf{b}$ has a solution for every \mathbf{b} .



Statement (10) implies Statement (11). Assume that T defined by $T(\mathbf{x}) = A\mathbf{x}$ is onto. For each i , let \mathbf{e}_i be the i th column of the $n \times n$ identity matrix I_n . That is, \mathbf{e}_i is the vector in \mathbb{R}^n with 1 in the i th component and 0 everywhere else. Since T is onto, for each i there is a vector \mathbf{c}_i such that $T(\mathbf{c}_i) = A\mathbf{c}_i = \mathbf{e}_i$. Let $C = [\mathbf{c}_1 \ \mathbf{c}_2 \ \cdots \ \mathbf{c}_n]$. Then

$$AC = A[\mathbf{c}_1 \ \mathbf{c}_2 \ \cdots \ \mathbf{c}_n] = [A\mathbf{c}_1 \ A\mathbf{c}_2 \ \cdots \ A\mathbf{c}_n] = [\mathbf{e}_1 \ \mathbf{e}_2 \ \cdots \ \mathbf{e}_n] = I_n.$$

Statement (11) implies Statement (12). Assume C is an $n \times n$ matrix so that $AC = I_n$. First we show that the matrix equation $C\mathbf{x} = \mathbf{0}$ has only the trivial solution. Suppose $C\mathbf{x} = \mathbf{0}$. Then multiplying both sides on the left by A gives us

$$A(C\mathbf{x}) = A\mathbf{0}.$$

Simplifying this equation using $AC = I_n$, we find $\mathbf{x} = \mathbf{0}$.

Since $C\mathbf{x} = \mathbf{0}$ has only the trivial solution, every column of C must be a pivot column. Since C is an $n \times n$ matrix, it follows that every row of C contains a pivot position. Thus, the matrix equation $C\mathbf{x} = \mathbf{b}$ is consistent and has a unique solution for every \mathbf{b} in \mathbb{R}^n . Let \mathbf{v}_i be the vector in \mathbb{R}^n satisfying $C\mathbf{v}_i = \mathbf{e}_i$ for each i between 1 and n and let $M = [\mathbf{v}_1 \ \mathbf{v}_2 \ \cdots \ \mathbf{v}_n]$. Then $CM = I_n$.

Now we show that $CA = I_n$. Since

$$AC = I_n$$

we can multiply both sides on the left by C to see that

$$C(AC) = CI_n.$$

Now we multiply both sides on the right by M and obtain

$$(C(AC))M = CM.$$

Using the associative property of matrix multiplication and the fact that $CM = I_n$ shows that

$$\begin{aligned} (CA)(CM) &= CM \\ CA &= I_n. \end{aligned}$$

Thus, if A and C are $n \times n$ matrices and $AC = I_n$, then $CA = I_n$. So we have proved our implication with $D = C$.

Statement (12) implies Statement (13). Assume that there is an $n \times n$ matrix D so that $DA = I_n$. Suppose $A\mathbf{x} = \mathbf{0}$. Then multiplying both sides by A on the left, we find that

$$\begin{aligned} D(A\mathbf{x}) &= D\mathbf{0} \\ (DA)\mathbf{x} &= \mathbf{0} \\ \mathbf{x} &= \mathbf{0}. \end{aligned}$$

So the equation $A\mathbf{x} = \mathbf{0}$ has only the trivial solution and 0 is not an eigenvalue for A .



Statement (13) implies Statement (14). If 0 is not an eigenvalue of A , then the equation $Ax = \mathbf{0}$ has only the trivial solution. Since statement 2 implies statement 11, there is an $n \times n$ matrix C such that $AC = I_n$. The proof that statement 11 implies statement 12 shows that $CA = I_n$ as well. So A is invertible. By taking the transpose of both sides of the equation $AA^{-1} = A^{-1}A = I_n$ (remembering $(AB)^T = B^T A^T$) we find

$$(A^{-1})^T A^T = A^T (A^{-1})^T = I_n^T = I_n.$$

Therefore, $(A^{-1})^T$ is the inverse of A^T by definition of the inverse.

Statement (14) implies Statement (1). Since statement (1) implies statement (14), we proved ‘If A is invertible, then A^T is invertible.’ Using this implication with A^T replaced by A , we find that ‘If A^T is invertible, then $(A^T)^T$ is invertible.’ But $(A^T)^T = A$, which proves that statement (14) implies statement (1).

This concludes our proof of the Invertible Matrix Theorem.

Examples

What follows are worked examples that use the concepts from this section.

Example 11.2. Let $M = \begin{bmatrix} 1 & 2 & 2 & 1 \\ 0 & 1 & 0 & 1 \\ 1 & 3 & 2 & 3 \\ 0 & 1 & 0 & 0 \end{bmatrix}$.

- (a) Without doing any calculations, is M invertible? Explain your response.
- (b) Is the equation $M\mathbf{x} = \mathbf{b}$ consistent for every \mathbf{b} in \mathbb{R}^4 ? Explain.
- (c) Is the equation $M\mathbf{x} = \mathbf{0}$ consistent? If so, how many solutions does this equation have? Explain.
- (d) Is it possible to find a 4×4 matrix P such that $PM = I_4$? Explain.

Example Solution.

- (a) The third column of M is twice the first, so the columns of M are not linearly independent. We conclude that M is not invertible.
- (b) The equation $M\mathbf{x} = \mathbf{b}$ is not consistent for every \mathbf{b} in \mathbb{R}^4 . If it was, then the columns of M would span \mathbb{R}^4 and, since there are exactly four columns, the columns of M would be a basis for \mathbb{R}^4 . Thus, M would have to be invertible, which it is not.
- (c) The homogeneous system is always consistent. Since the columns of M are linearly dependent, the equation $M\mathbf{x} = \mathbf{0}$ has infinitely many solutions.
- (d) It is not possible to find a 4×4 matrix P such that $PM = I_4$. Otherwise M would have to be invertible.



Example 11.3. Let M be an $n \times n$ matrix whose eigenvalues are all nonzero.

- Let $\mathbf{b} \in \mathbb{R}^n$. Is the equation $M\mathbf{x} = \mathbf{b}$ consistent? If yes, explain why and find all solutions in terms of M and \mathbf{b} . If no, explain why.
- Let S be the matrix transformation defined by $S(\mathbf{x}) = M\mathbf{x}$. Suppose $S(\mathbf{a}) = S(\mathbf{b})$ for some vectors \mathbf{a} and \mathbf{b} in \mathbb{R}^n . Must there be any relationship between \mathbf{a} and \mathbf{b} ? If yes, explain the relationship. If no, explain why.
- Let $\mathbf{m}_1, \mathbf{m}_2, \dots, \mathbf{m}_n$ be the columns of M . In how many ways can we write the zero vector as a linear combination of $\mathbf{m}_1, \mathbf{m}_2, \dots, \mathbf{m}_n$? Explain.

Example Solution.

- Since 0 is not an eigenvalue of M , we know that M is invertible. Therefore, the equation $M\mathbf{x} = \mathbf{b}$ has the unique solution $\mathbf{x} = M^{-1}\mathbf{b}$.
- The fact that M is invertible implies that S is one-to-one. So if $S(\mathbf{a}) = S(\mathbf{b})$, then it must be the case that $\mathbf{a} = \mathbf{b}$.
- Because M is invertible, the columns of M are linearly independent. Therefore, there is only the trivial solution to the equation

$$x_1\mathbf{m}_1 + x_2\mathbf{m}_2 + \cdots + x_n\mathbf{m}_n = \mathbf{0}.$$

Summary

- Two statements are equivalent if, whenever one of the statements is true, then the other must also be true.
- To efficiently prove that a string of statements are all equivalent, we can prove that each statement in the list implies the next statement, and that the last statement implies the first.
- The Invertible Matrix Theorem gives us many conditions that are equivalent to an $n \times n$ matrix being invertible. This theorem is important because it connects many of the concepts we have been studying.

Exercises

- (1) Consider the matrix $A = \begin{bmatrix} 1 & 2 & a \\ -1 & 1 & b \\ 1 & 1 & c \end{bmatrix}$. Use the Invertible Matrix Theorem to geometrically describe the vectors $\begin{bmatrix} a \\ b \\ c \end{bmatrix}$ which make A invertible without doing any calculations.



- (2) Suppose A is an invertible $n \times n$ matrix. Let T be the matrix transformation defined by $T(\mathbf{x}) = A\mathbf{x}$ for \mathbf{x} in \mathbb{R}^n . Show that the matrix transformation S defined by $S(\mathbf{x}) = A^{-1}\mathbf{x}$ is the inverse of the transformation T (i.e., S is the inverse function to T when the transformations are considered as functions).
- (3) Label each of the following statements as True or False. Provide justification for your response.
- True/False** If A^2 is invertible, then A is invertible.
 - True/False** If A and B are square matrices with AB invertible, then A and B are invertible.
 - True/False** If the columns of an $n \times n$ matrix A span \mathbb{R}^n , then the equation $A^{-1}\mathbf{x} = \mathbf{0}$ has a unique solution.
 - True/False** If the columns of A and columns of B form a basis of \mathbb{R}^n , then so do the columns of AB .
 - True/False** If the columns of a matrix A form a basis of \mathbb{R}^n , then so do the rows of A .
 - True/False** If the matrix transformation T defined by $T(\mathbf{x}) = A\mathbf{x}$ is one-to-one for an $n \times n$ matrix A , then the columns of A^{-1} are linearly independent.
 - True/False** If the columns of an $n \times n$ matrix A span \mathbb{R}^n , then so do the rows of A .
 - True/False** If there are two $n \times n$ matrices A and B such that $AB = I_n$, then the matrix transformation defined by $T(\mathbf{x}) = A^T\mathbf{x}$ is one-to-one.



Part III

The Vector Space \mathbb{R}^n

Section 12

The Structure of \mathbb{R}^n

Focus Questions

By the end of this section, you should be able to give precise and thorough answers to the questions listed below. You may want to keep these questions in mind to focus your thoughts as you complete the section.

- What properties make \mathbb{R}^n a vector space?
- What is a subspace of \mathbb{R}^n ?
- What properties do we need to verify to show that a set of vectors is a subspace of \mathbb{R}^n ? Why?
- What important structure does the span of a set of vectors in \mathbb{R}^n have?

Application: Connecting GDP and Consumption in Romania

It is common practice in the sciences to run experiments and collect data. Once data is collected it is necessary to find some way to analyze the data and predict future behavior from the data. One method is to find a curve that best “fits” the data, and one widely used method for curve fitting is called the *least squares* method.

For example, economists are often interested in *consumption*, which is the purchase of goods and services for use by households. In “A Statistical Analysis of GDP and Final Consumption Using Simple Linear Regression, the Case of Romania 1990-2010”,¹ the authors collect data and then use simple linear regression to compare GDP (gross domestic product) to consumption in Romania. The data they used is seen in Table 12.1, with a corresponding scatterplot of the data (with consumption as independent variable and GDP as dependent variable). The units for GDP

¹Bălăcescu, Aniela & Zaharia, Marian. (2012). A STATISTICAL ANALYSIS OF GDP AND FINAL CONSUMPTION USING SIMPLE LINEAR REGRESSION. THE CASE OF ROMANIA 1990?2010. Annals - Economy Series. 4. 26-31. Available from: https://www.researchgate.net/publication/227382939_A_STATISTICAL_ANALYSIS_OF_GDP_AND_FINAL_CONSUMPTION_USING_SIMPLE_LINEAR_REGRESSION_THE_CASE_OF_ROMANIA_1990-2010.

and consumption are millions of leu (the currency of Romania is the leu – on December 21, 2018, one leu was worth approximately \$0.25 U.S.) The authors conclude their paper with the following statement:

“However, we can appreciate that linear regression model describes the correlation between the value of gross domestic product and the value of final consumption and may be transcribed following form:

$$\text{PIB} = -3127.51 + 1.22 \text{ CF.}$$

Analysis of correlation between GDP and final consumption (private consumption and public consumption) will result in an increase of 1.22 units of monetary value of gross domestic product.

We can conclude that the Gross Domestic Product of our country is strongly influenced by the private and public consumption.”

Year	GDP	Consumption
1990	85.8	68.0
1991	220.4	167.3
1992	602.9	464.3
1993	2003.9	1523.6
1994	4977.3	3845.2
1995	7648.9	6257.7
1996	11384.2	9713.8
1997	25529.8	21972.2
1998	37055.1	33311.2
1999	55191.4	49311.9
2000	80984.6	69587.4
2001	117945.8	100731.7
2002	152017.0	127118.8
2003	197427.6	168818.7
2004	247368.0	211054.6
2005	288954.6	251038.1
2006	344650.6	294867.6
2007	416006.8	344937.0
2008	514700.0	420917.5
2009	498007.5	402246.0
2010	513640.8	405422.4

Table 12.1: GDP and consumption in Romania.

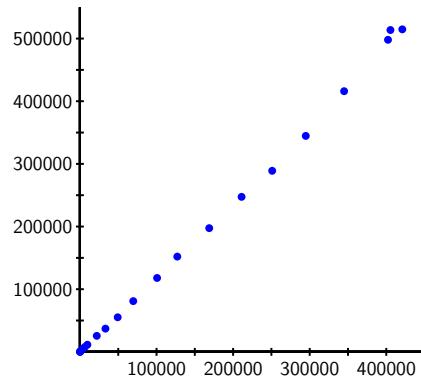


Figure 12.1: GDP and consumption data plot.

As we can see from the scatterplot, the relationship between the GDP and consumption is not exactly linear, but looks to be very close. To make correlations between GDP and consumption as the authors did, we need to understand how they determined their approximate linear relationship

between the variables. With a good approximation function we can then compare the variables, extrapolate from the data, and make predictions or interpolate and estimate between data points. For example, we could use our approximation function to predict, as the authors did, how changes in consumption affect GDP (or vice versa). Later in this section we will see how to find the least squares line to fit this data – the best linear approximation to the data. This involves finding a vector in a certain subspace of \mathbb{R}^2 that is closest to a given vector. Linear least squares approximation is a special case of a more general process that we will encounter in later sections where we learn how to project sets onto subspaces.

Introduction

The set \mathbb{R}^n with vector addition and scalar multiplication has a nice algebraic structure. These operations satisfy a number of properties, such as associativity and commutativity of vector addition, the existence of an additive identity and additive inverse, distribution of scalar multiplication over vector addition, and others. These properties make it easier to work with the whole space since we can express the vectors as linear combinations of basis vectors in a unique way. This algebraic structure makes \mathbb{R}^n a *vector space*.

There are many subsets of \mathbb{R}^n that have this same structure. These subsets are called *subspaces* of \mathbb{R}^n . These are the sets of vectors for which the addition of any two vectors is defined within the set, the scalar multiple of any vector by any scalar is defined within the set and the set contains the zero vector. One type of subset with this structure is the span of a set of vectors.

Recall that the span of a set of vectors $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k\}$ in \mathbb{R}^n is the set of all linear combinations of the vectors. For example, if $\mathbf{v}_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$ and $\mathbf{v}_2 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$, then a linear combination of these two vectors is of the form

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 = c_1 \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + c_2 \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} c_1 + c_2 \\ c_1 \\ c_2 \end{bmatrix}.$$

One linear combination can be obtained by letting $c_1 = 2, c_2 = -3$, which gives the vector $2\mathbf{v}_1 - 3\mathbf{v}_2 = \begin{bmatrix} -1 \\ -3 \\ 2 \end{bmatrix}$. All such linear combinations form the span of the vectors \mathbf{v}_1 and \mathbf{v}_2 . In this case, these vectors will form a plane through the origin in \mathbb{R}^3 .

Now we will investigate if the span of two vectors form a subspace, i.e. if it has the same structure as a vector space.

Preview Activity 12.1. Let \mathbf{w}_1 and \mathbf{w}_2 be two vectors in \mathbb{R}^n . Let $W = \text{Span}\{\mathbf{w}_1, \mathbf{w}_2\}$.

(1) For W to be a subspace of \mathbb{R}^n , the sum of any two vectors in W must also be in W .

- (a) Pick two specific examples of vectors \mathbf{u}, \mathbf{y} in W (keeping $\mathbf{w}_1, \mathbf{w}_2$ unknown/general vectors). For example, one specific \mathbf{u} would be $2\mathbf{w}_1 - 3\mathbf{w}_2$ as we used in the above example. Find the sum of \mathbf{u}, \mathbf{y} . Is the sum also in W ? Explain. (Hint: What does it mean for a vector to be in W ?)



- (b) Now let \mathbf{u} and \mathbf{y} be arbitrary vectors in W . Explain why $\mathbf{u} + \mathbf{y}$ is in W .
- (2) For W to be a subspace of \mathbb{R}^n , any scalar multiple of any vector in W must also be in W .
- Pick a specific example \mathbf{u} in W . Explain why $2\mathbf{u}, -3\mathbf{u}, \pi\mathbf{u}$ are all also in W .
 - Now let a be an arbitrary scalar and let \mathbf{u} be an arbitrary vector in W . Explain why the vector $a\mathbf{u}$ is in W .
- (3) For W to be a subspace of \mathbb{R}^n , the zero vector must also be in W . Explain why the zero vector is in W .
- (4) Does vector addition being commutative for vectors in \mathbb{R}^n imply that vector addition is also commutative for vectors in W ? Explain your reasoning.
- (5) Suppose we have an arbitrary \mathbf{u} in W . There is an additive inverse of \mathbf{u} in \mathbb{R}^n . In other words, there is a \mathbf{u}' such that $\mathbf{u} + \mathbf{u}' = \mathbf{0}$. Should this \mathbf{u}' be also in W ? If so, explain why. If not, give a counterexample.
- (6) Look at the other properties of vector addition and scalar multiplication of vectors in \mathbb{R}^n listed in Theorem 4.3 in Section 4. Which of these properties should also hold for vectors in W ?

Vector Spaces

The set of n -dimensional vectors with the vector addition and scalar multiplication satisfy many properties, such as addition being commutative and associative, existence of an additive identity, and others. The set \mathbb{R}^n with these properties is an example of a *vector space*, a general structure examples of which include many other algebraic structures as we will see later.

Definition 12.1. A set V on which an operation of addition and a multiplication by scalars is defined is a **vector space** if for all \mathbf{u}, \mathbf{v} , and \mathbf{w} in V and all scalars a and b :

- $\mathbf{u} + \mathbf{v}$ is an element of V (we say that V is *closed* under the addition in V),
- $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$ (we say that the addition in V is *commutative*),
- $(\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w})$ (we say that the addition in V is *associative*),
- there is a vector $\mathbf{0}$ in V so that $\mathbf{u} + \mathbf{0} = \mathbf{u}$ (we say that V contains an *additive identity* or *zero vector* $\mathbf{0}$),
- for each \mathbf{x} in V there is an element \mathbf{y} in V so that $\mathbf{x} + \mathbf{y} = \mathbf{0}$ (we say that V contains an *additive inverse* \mathbf{y} for each element \mathbf{x} in V),
- $a\mathbf{u}$ is an element of V (we say that V is *closed* under multiplication by scalars),
- $(a + b)\mathbf{u} = a\mathbf{u} + b\mathbf{u}$ (we say that *multiplication by scalars distributes over scalar addition*),
- $a(\mathbf{u} + \mathbf{v}) = a\mathbf{u} + a\mathbf{v}$ (we say that *multiplication by scalars distributes over addition in V*),



$$(9) \quad (ab)\mathbf{u} = a(b\mathbf{u}),$$

$$(10) \quad 1\mathbf{u} = \mathbf{u}.$$

Theorem 4.3 in Section 4 shows that \mathbb{R}^n is itself a vector space. As we will see, there are many other sets that have the same algebraic structure. By focusing on this structure and the properties of these operations, we can extend the theory of vectors we developed so far to a broad range of objects, making it easier to work with them. For example, we can consider linear combinations of functions or matrices, or define a basis for different types of sets of objects. Such algebraic tools provide us with new ways of looking at these sets of objects, including a geometric intuition when working with these sets. In this section, we will analyze subsets of \mathbb{R}^n which behave similar to \mathbb{R}^n algebraically. We will call such sets *subspaces*. In a later chapter we will encounter different kinds of sets that are also vector spaces.

Definition 12.2. A subset W of \mathbb{R}^n is a **subspace** of \mathbb{R}^n if W itself is a vector space using the same operations as in \mathbb{R}^n .

The following example illustrates the process for demonstrating that a subset of \mathbb{R}^n is a subspace of \mathbb{R}^n .

Example 12.3. There are many subsets of \mathbb{R}^n that are themselves vector spaces. Consider as an example the set W of vectors in \mathbb{R}^2 defined by

$$W = \left\{ \begin{bmatrix} x \\ 0 \end{bmatrix} \middle| x \text{ is a real number} \right\}.$$

In other words, W is the set of vectors in \mathbb{R}^2 whose second component is 0. To see that W is itself a vector space, we need to demonstrate that W satisfies all of the properties listed in Definition 12.1.

To prove the first property, we need to show that the sum of any two vectors in W is again in W . So we need to choose two arbitrary vectors in W . Let $\mathbf{u} = \begin{bmatrix} x \\ 0 \end{bmatrix}$ and $\mathbf{v} = \begin{bmatrix} y \\ 0 \end{bmatrix}$ be vectors in W . Note that

$$\mathbf{u} + \mathbf{v} = \begin{bmatrix} x \\ 0 \end{bmatrix} + \begin{bmatrix} y \\ 0 \end{bmatrix} = \begin{bmatrix} x+y \\ 0 \end{bmatrix}.$$

Since the second component of $\mathbf{u} + \mathbf{v}$ is 0, it follows that $\mathbf{u} + \mathbf{v}$ is in W . Thus, the set W is closed under addition.

For the second property, that addition is commutative in W , we can just use the fact that if \mathbf{u} and \mathbf{v} are in W , they are also vectors in \mathbb{R}^2 and $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$ is satisfied in \mathbb{R}^2 . So the property also holds in W .

A similar argument can be made for property (3).

Property (4) states the existence of the additive identity in W . Note that $\mathbf{0}$ is an additive identity in \mathbb{R}^2 and if it is also an element in W , then it will automatically be the additive identity of W . Since the zero vector can be written as $\mathbf{0} = \begin{bmatrix} x \\ 0 \end{bmatrix}$ with $x = 0$, $\mathbf{0}$ is in W . Thus, W satisfies property 4.

We will postpone property (5) for a bit since we can show that other properties imply property (5).



Property (6) is a closure property, just like property (1). We need to verify that any scalar multiple of any vector in W is again in W . Consider an arbitrary vector \mathbf{u} and an arbitrary scalar a . Now

$$a\mathbf{u} = a \begin{bmatrix} x \\ 0 \end{bmatrix} = \begin{bmatrix} ax \\ 0 \end{bmatrix}.$$

Since the vector $a\mathbf{u}$ has a 0 as its second component, we see that $a\mathbf{u}$ is in W . Thus, W is closed under scalar multiplication.

Properties (7), (8), (9) and (10) only depend on the operations of addition and multiplication by scalars in \mathbb{R}^2 . Since these properties depend on the operations and not the vectors, these properties will transfer to W .

We still have to justify property (5) though. Note that since $1 - 1 = 0$ in real numbers, by applying property (7) with $a = 1, b = -1$, we find that

$$\mathbf{0} = 0\mathbf{u} = (a + b)\mathbf{u} = a\mathbf{u} + b\mathbf{u} = \mathbf{u} + (-1)\mathbf{u}.$$

Therefore, $(-1)\mathbf{u}$ is an additive inverse for \mathbf{u} . Therefore, to show that the additive inverse of any \mathbf{u} in W is also in W , we simply note that any multiple of \mathbf{u} is also in W and hence $(-1)\mathbf{u}$ must also be in W .

Since W satisfies all of the properties of a vector space, W is a vector space. Any subset of \mathbb{R}^n that is itself a vector space using the same operations as in \mathbb{R}^n is called a *subspace* of \mathbb{R}^n .

Example 12.3 and our work Preview Activity 12.1 bring out some important ideas. When checking that a subset W of a vector space \mathbb{R}^n is also a vector space, we can use the fact that all of the properties of the operations in \mathbb{R}^n are transferred to any closed subset W . This implies that properties (2), (3), (7)-(10) are all automatically satisfied for W as well. Property (5) follows from the others. So we only need to check properties (1), (4) and (6). In fact, as we argued in the above example, property (4) also needs to be checked by simply checking that $\mathbf{0}$ of \mathbb{R}^n is in W . We summarize this result in the following theorem.

Theorem 12.4. *A subset W of \mathbb{R}^n is a subspace of \mathbb{R}^n if*

- (1) whenever \mathbf{u} and \mathbf{v} are in W it is also true that $\mathbf{u} + \mathbf{v}$ is in W (that is, W is **closed under addition**),
- (2) whenever \mathbf{u} is in W and a is a scalar it is also true that $a\mathbf{u}$ is in W (that is, W is **closed under scalar multiplication**),
- (3) $\mathbf{0}$ is in W .

The next activity provides some practice using Theorem 12.4.

Activity 12.1. Use Theorem 12.4 to answer the following questions. Justify your responses. For sets which lie inside \mathbb{R}^2 , sketch a pictorial representation of the set and explain why your picture confirms your answer.

- (a) Is the set $W = \left\{ \begin{bmatrix} x \\ y \end{bmatrix} \mid y = 2x \right\}$ a subspace of \mathbb{R}^2 ?



- (b) Is the set $W = \left\{ \begin{bmatrix} x \\ 0 \\ 1 \end{bmatrix} \middle| x \text{ is a scalar} \right\}$ a subspace of \mathbb{R}^3 ?
- (c) Is the set $W = \left\{ \begin{bmatrix} x \\ x+y \end{bmatrix} \middle| x, y \text{ are scalars} \right\}$ a subspace of \mathbb{R}^2 ?
- (d) Is the set $W = \left\{ \begin{bmatrix} x \\ y \end{bmatrix} \middle| y = 2x + 1 \right\}$ a subspace of \mathbb{R}^2 ?
- (e) Is the set $W = \left\{ \begin{bmatrix} x \\ y \end{bmatrix} \middle| y = x^2 \right\}$ a subspace of \mathbb{R}^2 ?
- (f) Is the set $W = \left\{ \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \right\}$ a subspace of \mathbb{R}^4 ?
- (g) Is the set $W = \left\{ \begin{bmatrix} x \\ y \\ z \end{bmatrix} \middle| x^2 + y^2 + z^2 \leq 1 \right\}$ a subspace of \mathbb{R}^3 ? Note that W is the unit sphere (a.k.a. unit ball) in \mathbb{R}^3 .
- (h) Is the set $W = \mathbb{R}^2$ a subspace of \mathbb{R}^3 ?

There are several important points that we can glean from Activity 12.1.

- A subspace is a vector space within a larger vector space, similar to a subset being a set within a larger set.
- The set containing the zero vector in \mathbb{R}^n is a subspace of \mathbb{R}^n , and it is the only finite subspace of \mathbb{R}^n .
- Every subspace of \mathbb{R}^n must contain the zero vector.
- No nonzero subspace is bounded – since a subspace must include all scalar multiples of its vectors, a subspace cannot be contained in a finite sphere or box.
- Since vectors in \mathbb{R}^k have k components, vectors in \mathbb{R}^k are not contained in \mathbb{R}^n when $n \neq k$. However, if $n > k$, then we can think of \mathbb{R}^n as containing a *copy* (what we call an isomorphic image) of \mathbb{R}^k as the set of vectors with zeros as the last $n - k$ components.

The Subspace Spanned by a Set of Vectors

One of the most convenient ways to represent a subspace of \mathbb{R}^n is as the span of a set of vectors. In Preview Activity 12.1 we saw that the span of two vectors is a subspace of \mathbb{R}^n . In the next theorem we verify this result for the span of an arbitrary number of vectors, extending the ideas you used in Preview Activity 12.1. Expressing a set of vectors as the span of some number of vectors is a quick



way of justifying that this set is a subspace and it also provides us a geometric intuition for the set of vectors.

Theorem 12.5. *Let $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k$ be vectors in \mathbb{R}^n . Then $\text{Span}\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k\}$ is a subspace of \mathbb{R}^n .*

Proof. Let $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k$ be vectors in \mathbb{R}^n . Let $W = \text{Span}\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k\}$. To show that W is a subspace of \mathbb{R}^n we need to show that W is closed under addition and multiplication by scalars and that $\mathbf{0}$ is in W .

First we show that W is closed under addition. Let \mathbf{u} and \mathbf{w} be vectors in W . This means that \mathbf{u} and \mathbf{w} are linear combinations of $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k$. So there are scalars a_1, a_2, \dots, a_k and b_1, b_2, \dots, b_k so that

$$\mathbf{u} = a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \cdots + a_k\mathbf{v}_k \quad \text{and} \quad \mathbf{w} = b_1\mathbf{v}_1 + b_2\mathbf{v}_2 + \cdots + b_k\mathbf{v}_k.$$

To demonstrate that $\mathbf{u} + \mathbf{w}$ is in W , we need to show that $\mathbf{u} + \mathbf{w}$ is a linear combination of $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k$. Using the properties of vector addition and scalar multiplication, we find

$$\begin{aligned} \mathbf{u} + \mathbf{w} &= (a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \cdots + a_k\mathbf{v}_k) + (b_1\mathbf{v}_1 + b_2\mathbf{v}_2 + \cdots + b_k\mathbf{v}_k) \\ &= (a_1 + b_1)\mathbf{v}_1 + (a_2 + b_2)\mathbf{v}_2 + \cdots + (a_k + b_k)\mathbf{v}_k. \end{aligned}$$

Thus $\mathbf{u} + \mathbf{w}$ is a linear combination of $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k$ and W is closed under vector addition.

Next we show that W is closed under scalar multiplication. Let \mathbf{u} be in W and c be a scalar. Then

$$c\mathbf{u} = c(a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \cdots + a_k\mathbf{v}_k) = (ca_1)\mathbf{v}_1 + (ca_2)\mathbf{v}_2 + \cdots + (ca_k)\mathbf{v}_k$$

and $c\mathbf{u}$ is a linear combination of $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k$ and W is closed under multiplication by scalars.

Finally, we show that $\mathbf{0}$ is in W . Since

$$\mathbf{0} = 0\mathbf{v}_1 + 0\mathbf{v}_2 + \cdots + 0\mathbf{v}_k,$$

$\mathbf{0}$ is in W .

Since W satisfies all of the properties of a subspace as given in definition of a subspace, we conclude that W is a subspace of \mathbb{R}^n . ■

The subspace $W = \text{Span}\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k\}$ is called the *subspace of \mathbb{R}^n spanned by $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k$* . We also use the phrase “subspace generated by $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k$ ” since the vectors $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k$ are the building blocks of all vectors in W .

Activity 12.2.

- (a) Describe geometrically as best as you can the subspaces of \mathbb{R}^3 spanned by the following sets of vectors.

$$\left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \right\}$$

$$\left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \right\}$$

- (b) Express the following set of vectors as the span of some vectors to show that this set is a subspace. Can you give a geometric description of the set?

$$W = \left\{ \begin{bmatrix} 2x + y - z \\ y \\ z \\ -x + 3z \end{bmatrix} : x, y, z \text{ real numbers} \right\}$$

One additional conclusion we can draw from Activities 12.1 and 12.2 is that subspaces of \mathbb{R}^n are made up of “flat” subsets. The span of a single nonzero vector is a line (which is flat), and the span of a set of two distinct nonzero vectors is a plane (which is also flat). So subspaces of \mathbb{R}^n are linear (or “flat”) subsets of \mathbb{R}^n . That is why we can recognize that the non-flat parabola in Activity 12.1 is not a subspace of \mathbb{R}^2 .

Examples

What follows are worked examples that use the concepts from this section.

Example 12.6. Let $W = \left\{ \begin{bmatrix} 2r + s + t \\ r + t \\ r + s \end{bmatrix} : r, s, t \in \mathbb{R} \right\}$.

- (a) Show that W is a subspace of \mathbb{R}^3 .
- (b) Describe in detail the geometry of the subspace W (e.g., is it a line, a union of lines, a plane, a union of planes, etc.)

Example Solution.

- (a) Every vector in W has the form

$$\begin{bmatrix} 2r + s + t \\ r + t \\ r + s \end{bmatrix} = r \begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix} + s \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} + t \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$$

for some real numbers r , s , and t . Thus,

$$W = \text{Span} \left\{ \begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} \right\}.$$

As a span of a set of vectors, we know that W is a subspace of \mathbb{R}^4 .

- (b) Let $\mathbf{v}_1 = \begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix}$, $\mathbf{v}_2 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$, and $\mathbf{v}_3 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$. The reduced row echelon form of $[\mathbf{v}_1 \mathbf{v}_2 \mathbf{v}_3]$ is $\begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{bmatrix}$. The pivot columns of $[\mathbf{v}_1 \mathbf{v}_2 \mathbf{v}_3]$ form a linearly independent



set with the same span as $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$, So $W = \text{Span}\{\mathbf{v}_1, \mathbf{v}_2\}$ and W forms the plane in \mathbb{R}^3 through the origin and the points $(2, 1, 1)$ and $(1, 0, 1)$.

Example 12.7.

(a) Let $X = \text{Span}\left\{\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}\right\}$ and let $Y = \text{Span}\left\{\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}\right\}$. That is, X is the x -axis and Y the y -axis in three-space. Let

$$X + Y = \{\mathbf{x} + \mathbf{y} : \mathbf{x} \in X \text{ and } \mathbf{y} \in Y\}.$$

i. Is $\begin{bmatrix} 2 \\ 3 \\ 0 \end{bmatrix}$ in $X + Y$? Justify your answer.

ii. Is $\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ in $X + Y$? Justify your answer.

iii. Assume that $X + Y$ is a subspace of \mathbb{R}^3 . Describe in detail the geometry of this subspace.

(b) Now let W_1 and W_2 be arbitrary subspaces of \mathbb{R}^n for some positive integer n . Let

$$W_1 + W_2 = \{\mathbf{w}_1 + \mathbf{w}_2 : \mathbf{w}_1 \in W_1 \text{ and } \mathbf{w}_2 \in W_2\}.$$

Show that $W_1 + W_2$ is a subspace of \mathbb{R}^n . The set $W_1 + W_2$ is called the *sum* of the subspaces W_1 and W_2 .

Example Solution.

(a) We let $X = \text{Span}\left\{\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}\right\}$ and $Y = \text{Span}\left\{\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}\right\}$. T

i. Let $\mathbf{w} = \begin{bmatrix} 2 \\ 3 \\ 0 \end{bmatrix}$, $\mathbf{x} = 2 \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$, and $\mathbf{y} = 3 \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$. Since $\mathbf{w} = \mathbf{x} + \mathbf{y}$ with $\mathbf{x} \in X$ and $\mathbf{y} \in Y$ we conclude that $\mathbf{w} \in X + Y$.

ii. Every vector in X has the form $a\mathbf{e}_1$ for some scalar a (where $\mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$, and every

vector in Y has the form $b\mathbf{e}_2$ for some scalar b (where $\mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$). So every vector

in $X + Y$ is of the form $a\mathbf{e}_1 + b\mathbf{e}_2 = \begin{bmatrix} a \\ b \\ 0 \end{bmatrix}$. Since the vector $\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ does not have

a 0 in the third component, we conclude that in $\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ is not in $X + Y$.



iii. As we just argued, every vector in $X + Y$ has the form $a\mathbf{e}_1 + b\mathbf{e}_2$. So $X + Y = \text{Span}\{\mathbf{e}_1, \mathbf{e}_2\}$, which is the xy -plane in \mathbb{R}^3 .

- (b) To see why the set $W_1 + W_2$ is a subspace of \mathbb{R}^3 , suppose that \mathbf{x} and \mathbf{y} are in $W_1 + W_2$. Then $\mathbf{x} = \mathbf{u}_1 + \mathbf{u}_2$ and $\mathbf{y} = \mathbf{z}_1 + \mathbf{z}_2$ for some $\mathbf{u}_1, \mathbf{z}_1$ in W_1 and some $\mathbf{u}_2, \mathbf{z}_2$ in W_2 . Then

$$\mathbf{x} + \mathbf{y} = (\mathbf{u}_1 + \mathbf{u}_2) + (\mathbf{z}_1 + \mathbf{z}_2) = (\mathbf{u}_1 + \mathbf{z}_1) + (\mathbf{u}_2 + \mathbf{z}_2).$$

Since W_1 is a subspace of \mathbb{R}^3 it follows that $\mathbf{u}_1 + \mathbf{z}_1 \in W_1$. Similarly, $\mathbf{u}_2 + \mathbf{z}_2 \in W_2$. This makes $\mathbf{x} + \mathbf{y}$ an element of $W_1 + W_2$.

Also, suppose that a is a scalar. Then

$$a\mathbf{x} = a(\mathbf{u}_1 + \mathbf{u}_2) = a\mathbf{u}_1 + a\mathbf{u}_2.$$

Since W_1 is a subspace of \mathbb{R}^3 it follows that $a\mathbf{u}_1 \in W_1$. Similarly, $a\mathbf{u}_2 \in W_2$. This makes $a\mathbf{x}$ an element of $W_1 + W_2$.

Finally, since $\mathbf{0}$ is in both W_1 and W_2 , and $\mathbf{0} = \mathbf{0} + \mathbf{0}$, it follows that $\mathbf{0}$ is an element of $W_1 + W_2$. We conclude that $W_1 + W_2$ is a subspace of \mathbb{R}^3 .

Summary

- A vector space is a set V with operations of addition and scalar multiplication defined on V such that for all \mathbf{u}, \mathbf{v} , and \mathbf{w} in V and all scalars a and b :
 - (1) $\mathbf{u} + \mathbf{v}$ is an element of V (we say that V is *closed* under the addition in V),
 - (2) $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$ (we say that the addition in V is *commutative*),
 - (3) $(\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w})$ (we say that the addition in V is *associative*),
 - (4) there is a vector $\mathbf{0}$ in V so that $\mathbf{u} + \mathbf{0} = \mathbf{u}$ (we say that V contains an *additive identity* or *zero vector* $\mathbf{0}$),
 - (5) for each \mathbf{x} in V there is an element \mathbf{y} in V so that $\mathbf{x} + \mathbf{y} = \mathbf{0}$ (we say that V contains an *additive inverse* \mathbf{y} for each element \mathbf{x} in V),
 - (6) $a\mathbf{u}$ is an element of V (we say that V is *closed* under multiplication by scalars),
 - (7) $(a + b)\mathbf{u} = a\mathbf{u} + b\mathbf{u}$ (we say that *multiplication by scalars distributes over scalar addition*),
 - (8) $a(\mathbf{u} + \mathbf{v}) = a\mathbf{u} + a\mathbf{v}$ (we say that *multiplication by scalars distributes over addition in V*),
 - (9) $(ab)\mathbf{u} = a(b\mathbf{u})$,
 - (10) $1\mathbf{u} = \mathbf{u}$.
- For every n , \mathbb{R}^n is a vector space.
- A subset W of \mathbb{R}^n is a subspace of \mathbb{R}^n if W is a vector space using the same operations as in \mathbb{R}^n .
- To show that a subset W of \mathbb{R}^n is a subspace of \mathbb{R}^n , we need to prove the following:



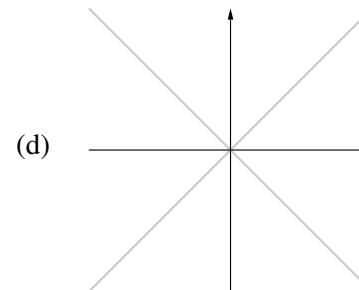
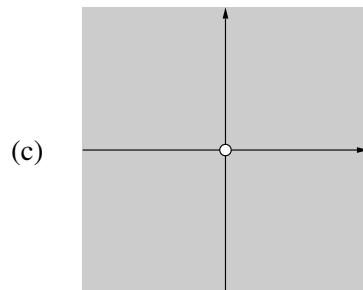
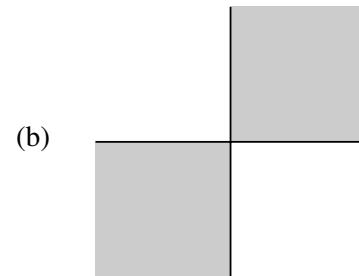
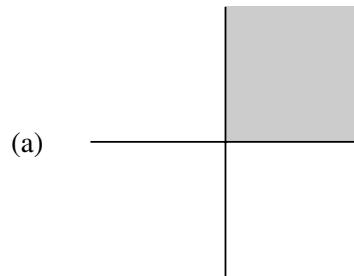
- (1) $\mathbf{u} + \mathbf{v}$ is in W whenever \mathbf{u} and \mathbf{v} are in W (when this property is satisfied we say that W is *closed* under addition),
- (2) $a\mathbf{u}$ is in W whenever a is a scalar and \mathbf{u} is in W (when this property is satisfied we say that W is *closed* under multiplication by scalars),
- (3) $\mathbf{0}$ is in W .

The remaining properties of a vector space are properties of the operation, and as long as we use the same operations as in \mathbb{R}^n , the operation properties follow the operations.

- The span of any set of vectors in \mathbb{R}^n is a subspace of \mathbb{R}^n .

Exercises

- (1) Each of the following regions or graphs determines a subset W of \mathbb{R}^2 . For each region, discuss each of the subspace properties of Theorem 12.4 and explain with justification if the set W satisfies each property or not.



- (2) Determine which of the following sets W is a subspace of \mathbb{R}^n for the indicated value of n . Justify your answer.

(a) $W = \{[x \ 0]^T : x \text{ is a scalar}\}$



- (b) $W = \{[2x + y \ x - y \ x + y]^T : x, y \text{ are scalars}\}$
 (c) $W = \{[x + 1 \ x - 1]^T : x \text{ is a scalar}\}$
 (d) $W = \{[xy \ xz \ yz]^T : x, y, z \text{ are scalars}\}$
- (3) Find a subset of \mathbb{R}^2 that is closed under addition and scalar multiplication, but that does not contain the zero vector, or explain why no such subset exists.
- (4) Let \mathbf{v} be a vector in \mathbb{R}^2 . What is the smallest subspace of \mathbb{R}^2 that contains \mathbf{v} ? Explain. Describe this space geometrically.
- (5) What is the smallest subspace of \mathbb{R}^2 containing the first quadrant? Justify your answer.
- (6) Let \mathbf{u} , \mathbf{v} , and \mathbf{w} be vectors in \mathbb{R}^3 with $\mathbf{w} = \mathbf{u} + \mathbf{v}$. Let $W_1 = \text{Span}\{\mathbf{u}, \mathbf{v}\}$ and $W_2 = \text{Span}\{\mathbf{u}, \mathbf{v}, \mathbf{w}\}$.
- (a) If \mathbf{x} is in W_1 , must \mathbf{x} be in W_2 ? Explain.
 - (b) If \mathbf{y} is in W_2 , must \mathbf{y} be in W_1 ? Explain.
 - (c) What is the relationship between $\text{Span}\{\mathbf{u}, \mathbf{v}\}$ and $\text{Span}\{\mathbf{u}, \mathbf{v}, \mathbf{w}\}$? Be specific.
- (7) Let m and n be positive integers, and let \mathbf{v} be in \mathbb{R}^n . Let $W = \{A\mathbf{v} : A \in \mathcal{M}_{m \times n}\}$.
- (a) As an example, let $\mathbf{v} = [2 \ 1]^T$ in \mathbb{R}^2 with $W = \{A\mathbf{v} : A \in \mathcal{M}_{2 \times 2}\}$.
 - i. Show that the vector $[2 \ 1]^T$ is in W by finding a matrix A that places $[2 \ 1]^T$ in W .
 - ii. Show that the vector $[4 \ 2]^T$ is in W by finding a matrix A that places $[4 \ 2]^T$ in W .
 - iii. Show that the vector $[6 \ -1]^T$ is in W by finding a matrix A that places $[6 \ -1]^T$ in W .
 - iv. Show that $W = \mathbb{R}^2$.
 - (b) Show that, regardless of the vector \mathbf{v} selected, W is a subspace of \mathbb{R}^m .
 - (c) Characterize all of the possibilities for what the subspace W can be. (Hint: There is more than one possibility.)
- (8) Let S_1 and S_2 be subsets of \mathbb{R}^3 such that $\text{Span } S_1 = \text{Span } S_2$. Must it be the case that S_1 and S_2 contain at least one vector in common? Justify your answer.
- (9) Assume W_1 and W_2 are two subspaces of \mathbb{R}^n . Is $W_1 \cap W_2$ also a subspace of \mathbb{R}^n ? Is $W_1 \cup W_2$ also a subspace of \mathbb{R}^n ? Justify your answer. (Note: The notation $W_1 \cap W_2$ refers to the vectors common to both W_1, W_2 , while the notation $W_1 \cup W_2$ refers to the vectors that are in at least one of W_1, W_2 .)
- (10) Determine whether the plane defined by the equation $5x + 3y - 2z = 0$ is a subspace in \mathbb{R}^3 .
- (11) If W is a subspace of \mathbb{R}^n and \mathbf{u} is a vector in \mathbb{R}^n not in W , determine whether

$$\mathbf{u} + W = \{\mathbf{u} + \mathbf{v} : \mathbf{v} \text{ is a vector in } W\}$$

is a subspace of \mathbb{R}^n .



- (12) Two students are talking about examples of subspaces.

Student 1: The x -axis in \mathbb{R}^2 is a subspace. It is generated by the vector $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$.

Student 2: Similarly \mathbb{R}^2 is a subspace of \mathbb{R}^3 .

Student 1: I'm not sure if that will work. Can we fit \mathbb{R}^2 inside \mathbb{R}^3 ? Don't we need W to be a subset of \mathbb{R}^3 if it is a subspace of \mathbb{R}^3 ?

Student 2: Of course we can fit \mathbb{R}^2 inside \mathbb{R}^3 . We can think of \mathbb{R}^2 as vectors $\begin{bmatrix} a \\ b \\ 0 \end{bmatrix}$. That's the xy -plane.

Student 1: I don't know. The vector $\begin{bmatrix} a \\ b \\ 0 \end{bmatrix}$ is not exactly same as $\begin{bmatrix} a \\ b \end{bmatrix}$.

Student 2: Well, \mathbb{R}^2 is a plane and so is the xy -plane. So they must be equal, shouldn't they?

Student 1: But there are infinitely many planes in \mathbb{R}^3 . They can't all be equal to \mathbb{R}^2 . They all "look like" \mathbb{R}^2 but I don't think we can say they are equal.

Which student is correct? Is \mathbb{R}^2 a subspace of \mathbb{R}^3 , or not? Justify your answer.

- (13) Label each of the following statements as True or False. Provide justification for your response.
- True/False** Any line in \mathbb{R}^n is a subspace in \mathbb{R}^n .
 - True/False** Any line through the origin in \mathbb{R}^n is a subspace in \mathbb{R}^n .
 - True/False** Any plane through the origin in \mathbb{R}^n is a subspace in \mathbb{R}^n .
 - True/False** In \mathbb{R}^4 , the points satisfying $xy = 2t + z$ form a subspace.
 - True/False** In \mathbb{R}^4 , the points satisfying $x + 3y = 2z$ form a subspace.
 - True/False** Any two nonzero vectors generate a plane subspace in \mathbb{R}^3 .
 - True/False** The space \mathbb{R}^2 is a subspace of \mathbb{R}^3 .

- True/False** If W is a subspace of \mathbb{R}^n and \mathbf{u} is in W , then the line through the origin and \mathbf{u} is in W .
- True/False** There are four types of subspaces in \mathbb{R}^3 : 0, line through origin, plane through origin and the whole space \mathbb{R}^3 .
- True/False** There are four types of subspaces in \mathbb{R}^4 : 0, line through origin, plane through origin and the whole space \mathbb{R}^4 .

(k) **True/False** The vectors $\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$, $\begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}$ and $\begin{bmatrix} 2 \\ 3 \\ 2 \end{bmatrix}$ form a subspace in \mathbb{R}^3 .

(l) **True/False** The vectors $\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ and $\begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}$ form a basis of a subspace in \mathbb{R}^3 .

Project: Least Squares Linear Approximation

We return to the problem of finding the least squares line to fit the GDP-consumption data. We will start our work in a more general setting, determining the method for fitting a linear function to fit any data set, like the GDP-consumption data, in the least squares sense. Then we will apply our result to the GDP-consumption data.

Project Activity 12.1. Suppose we want to fit a linear function $p(x) = mx + b$ to our data. For the sake of our argument, let us assume the general case where we have n data points labeled as $(x_1, y_1), (x_2, y_2), (x_3, y_3), \dots, (x_n, y_n)$. (In the GDP-consumption data $n = 21$.) In the unlikely event that the graph of $p(x)$ actually passes through these data points, then we would have the system of equations

$$\begin{aligned} y_1 &= b + mx_1 \\ y_2 &= b + mx_2 \\ y_3 &= b + mx_3 \\ &\vdots && \vdots \\ y_n &= b + mx_n \end{aligned} \tag{12.1}$$

in the unknowns b and m .

- (a) As a small example to illustrate, write the system (12.1) using the three points $(x_1, y_1) = (1, 2)$, $(x_2, y_2) = (3, 4)$, and $(x_3, y_3) = (5, 6)$. Identify the unknowns and then write this system in the form $M\mathbf{a} = \mathbf{y}$. Explicitly identify the matrix M and the vectors \mathbf{a} and \mathbf{y} .
- (b) Identify the specific matrix M and the specific vectors \mathbf{a} and \mathbf{y} using the data in Table 12.1. Explain why the system $M\mathbf{a} = \mathbf{y}$ is inconsistent. (Remember, we are treating consumption as the independent variable and GDP as the dependent variable.) What does the result tell us about the data?

Project Activity 12.1 shows that the GDP-consumption data does not lie on a line. So instead of attempting to find coefficients b and m that give a solution to this system, which may be impossible, we instead look for a vector \mathbf{a}^* that provides us with something that is “close” to a solution.

If we could find b and m that give a solution to the system $M\mathbf{a} = \mathbf{y}$, then $M\mathbf{a} - \mathbf{y}$ would be zero. If we can't make $M\mathbf{a} - \mathbf{y}$ exactly equal to the vector $\mathbf{0}$, we could instead try to minimize $M\mathbf{a} - \mathbf{y}$ in some way. One way is to minimize the length $\|M\mathbf{a} - \mathbf{y}\|$ of the vector $M\mathbf{a} - \mathbf{y}$.

If we minimize the quantity $\|M\mathbf{a} - \mathbf{y}\|$, then we will have minimized a function given by a sum of squares. That is, $\|M\mathbf{a} - \mathbf{y}\|$ is calculated to be

$$\sqrt{(b + mx_1 - y_1)^2 + (b + mx_2 - y_2)^2 + \dots + (b + mx_n - y_n)^2}. \tag{12.2}$$



This is why the method we will derive is called the method of least squares. This method provides us with a vector “solution” in a subspace that is related to M . We can visualize $\|M\mathbf{a} - \mathbf{y}\|$ as in Figure 12.2. In this figure the data points are shown along with a linear approximation (not the best for illustrative purposes). The lengths of the vertical line segments are the summands $(b + mx_i - y_i)$ in (12.2). So we are trying to minimize the sum of the squares of these line segments.

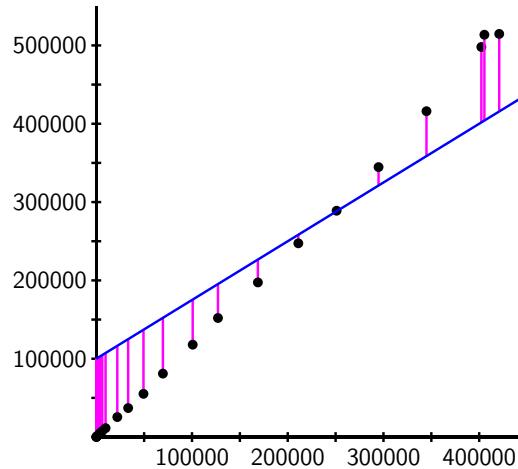


Figure 12.2: Error in the linear approximation.

Suppose that \mathbf{a}^* minimizes $\|M\mathbf{a} - \mathbf{y}\|$. Then the vector $M\mathbf{a}^*$ is the vector that is closest to \mathbf{y} of all of the vectors of the form $M\mathbf{x}$. The fact that the vectors of the form $M\mathbf{x}$ make a subspace will be useful in what follows. We verify that fact in the next project activity.

Project Activity 12.2. Let A be an arbitrary $m \times k$ matrix. Explain why the set $C = \{A\mathbf{x} : \mathbf{x} \in \mathbb{R}^k\}$ is a subspace of \mathbb{R}^m .

Project Activity 12.2 shows us that even though the GDP-consumption system $M\mathbf{a} = \mathbf{y}$ does not have a solution, we can find a vector that is close to a solution in the subspace $\{M\mathbf{x} : \mathbf{x} \in \mathbb{R}^2\}$. That is, find a vector \mathbf{a}^* in \mathbb{R}^2 such that $M\mathbf{a}^*$ is as close (in the least squares sense) to \mathbf{y} as we can get. In other words, the error $\|M\mathbf{a}^* - \mathbf{y}\|$ is as small as possible. In the following activity we see how to find \mathbf{a}^* .

Project Activity 12.3. Let

$$S = \sqrt{(b + mx_1 - y_1)^2 + (b + mx_2 - y_2)^2 + \cdots + (b + mx_n - y_n)^2},$$

the quantity we want to minimize. The variables in S are m and b , so we can think of S as a function of the two independent variables m and b . The square root makes calculations more complicated, so it is helpful to notice that S will be a minimum when S^2 is a minimum. Since S^2 is also function of the two variables b and m , the minimum value of S^2 will occur when the partial derivatives of S^2 with respect to b and m are both 0 (if you haven’t yet taken a multivariable calculus course, you

can just assume that this is correct). This yields the equations

$$0 = \sum_{i=1}^n (mx_i + b - y_i) x_i \quad (12.3)$$

$$0 = \sum_{i=1}^n (mx_i + b - y_i). \quad (12.4)$$

In this activity we solve equations (12.3) and (12.4) for the unknowns b and m . (Do this in a general setting without using specific values for the x_i and y_i .)

- (a) Let $r = \sum_{i=1}^n x_i^2$, $s = \sum_{i=1}^n x_i$, $t = \sum_{i=1}^n y_i$, and $u = \sum_{i=1}^n x_i y_i$. Show that the equations (12.3) and (12.4) can be written in the form

$$\begin{aligned} 0 &= bs + mr - u \\ 0 &= bn + ms - t. \end{aligned}$$

Note that this is a system of two linear equations in the unknowns b and m .

- (b) Write the system from part (a) in matrix form $A\mathbf{x} = \mathbf{b}$. Then use techniques from linear algebra to solve the linear system to show that

$$b = \frac{tr - us}{nr - s^2} = \frac{(\sum_{i=1}^n y_i)(\sum_{i=1}^n x_i^2) - (\sum_{i=1}^n x_i)(\sum_{i=1}^n x_i y_i)}{n(\sum_{i=1}^n x_i^2) - (\sum_{i=1}^n x_i)^2} \quad (12.5)$$

and

$$m = \frac{nu - ts}{nr - s^2} = \frac{n(\sum_{i=1}^n x_i y_i) - (\sum_{i=1}^n x_i)(\sum_{i=1}^n y_i)}{n(\sum_{i=1}^n x_i^2) - (\sum_{i=1}^n x_i)^2}. \quad (12.6)$$

Project Activity 12.4. Use the formulas (12.5) and (12.6) to find the values of b and m for the regression line to fit the GDP-consumption data in Table 12.1. You may use the fact that the sum of the GDP data is 3.5164030×10^6 , the sum of the consumption data is 2.9233750×10^6 , the sum of the squares of the consumption data is $8.806564894 \times 10^{11}$, and the sum of the products of the GDP and consumption data is $1.069946378 \times 10^{12}$. Compare to the results the authors obtained in the paper “A Statistical Analysis of GDP and Final Consumption Using Simple Linear Regression, the Case of Romania 1990-2010”.

Section 13

The Null Space and Column Space of a Matrix

Focus Questions

By the end of this section, you should be able to give precise and thorough answers to the questions listed below. You may want to keep these questions in mind to focus your thoughts as you complete the section.

- What is the null space of a matrix?
- What is the column space of a matrix?
- What important structure do the null space and column space of a matrix have?
- What is the kernel of a matrix transformation?
- How is the kernel of a matrix transformation T defined by $T(\mathbf{x}) = A\mathbf{x}$ related to the null space of A ?
- What is the range of a matrix transformation?
- How is the range of a matrix transformation T defined by $T(\mathbf{x}) = A\mathbf{x}$ related to the column space of A ?
- How do we find a basis for $\text{Nul } A$?
- How do we find a basis for $\text{Col } A$?

Application: The Lights Out Game

Lights Out (LO) is a commercial game released by Tiger Toys in 1995 (later bought out by Hasbro). The game consists of a 5×5 grid in which each square is either lit or unlit. Pressing a square changes the status of the square itself and all the squares to the left, right, up, or down. The player's job

is to turn all the lights out. You can play a sample game at <https://www.geogebra.org/m/wcmctaht>. There is a method to solve any solvable Lights Out game that can be uncovered through linear algebra that we will uncover later in this section. Column spaces and null spaces play important roles in this method.

Introduction

Recall that a subspace of \mathbb{R}^n is a subset of \mathbb{R}^n which is a vector space in itself. More specifically, a subset W or \mathbb{R}^n is a subspace of \mathbb{R}^n if

- (1) whenever \mathbf{u} and \mathbf{v} are in W it is also true that $\mathbf{u} + \mathbf{v}$ is in W (that is, W is **closed** under addition),
- (2) whenever \mathbf{u} is in W and a is a scalar it is also true that $a\mathbf{u}$ is in W (that is, W is **closed** under scalar multiplication),
- (3) $\mathbf{0}$ is in W .

Given a matrix A , there are several subspaces that are connected to A . Two specific such subspaces are the null space of A and the column space of A . We will see that these subspaces provide answers to the big questions we have been considering since the beginning of the semester, such as “Do columns of A span \mathbb{R}^m ?” “Are the columns of A linearly independent?” “Is the transformation T defined by matrix multiplication by A one-to-one?” “Is the transformation T onto?”

In this preview activity, we start examining the *null space*.

Preview Activity 13.1.

(1) Let $A = \begin{bmatrix} 2 & 1 & 3 \\ 1 & 1 & 4 \end{bmatrix}$.

- (a) Find the general solution to the homogeneous equation $A\mathbf{x} = \mathbf{0}$. Write your solutions in parametric vector form. (Recall that the parametric vector form expresses the solutions to an equation as linear combinations of vectors with free variables as the

weights. An example would be $x_3 \begin{bmatrix} 1 \\ 0 \\ -1 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} -2 \\ 1 \\ 0 \\ 1 \end{bmatrix}$.)

- (b) Find two specific solutions \mathbf{x}_1 and \mathbf{x}_2 to the homogeneous equation $A\mathbf{x} = \mathbf{0}$. Is $\mathbf{x}_1 + \mathbf{x}_2$ a solution to $A\mathbf{x} = \mathbf{0}$? Explain.

- (c) Is $3\mathbf{x}_1$ a solution to $A\mathbf{x} = \mathbf{0}$? Explain.

- (d) Is $\mathbf{0}$ a solution to $A\mathbf{x} = \mathbf{0}$?

- (e) What does the above seem to indicate about the set of solutions to the homogeneous system $A\mathbf{x} = \mathbf{0}$?

- (2) Let A be an $m \times n$ matrix. As problem 1 implies, the set of solutions to a homogeneous matrix-vector equation $A\mathbf{x} = \mathbf{0}$ appears to be a subspace. We give a special name to this set.



Definition 13.1. The **null space** of an $m \times n$ matrix A is the set of all solutions to $A\mathbf{x} = \mathbf{0}$.

We denote the null space of a matrix A as $\text{Nul } A$. In set notation we write

$$\text{Nul } A = \{\mathbf{x} : A\mathbf{x} = \mathbf{0}\}.$$

Note that since $A\mathbf{x} = \mathbf{0}$ corresponds to a homogeneous system of linear equations, $\text{Nul } A$ also represents the solution set of a homogeneous system.

Let $A = \begin{bmatrix} 2 & 1 & 3 & 0 \\ 1 & 1 & 4 & 1 \end{bmatrix}$. Find all vectors in $\text{Nul } A$.

- (3) So far we considered specific examples of null spaces. But what are the properties of a null space in general? Let A be an *arbitrary* $m \times n$ matrix.
- (a) The null space of an $m \times n$ matrix is a subset of \mathbb{R}^k for some integer k . What is k ?
 - (b) Now suppose \mathbf{u} and \mathbf{v} are two vectors in $\text{Nul } A$. By definition, that means $A\mathbf{u} = \mathbf{0}$, $A\mathbf{v} = \mathbf{0}$. Use properties of the matrix-vector product to show that $\mathbf{u} + \mathbf{v}$ is also in $\text{Nul } A$.
 - (c) Now suppose \mathbf{u} is a vector in $\text{Nul } A$ and a is a scalar. Explain why $a\mathbf{u}$ is also in $\text{Nul } A$.
 - (d) Explain why $\text{Nul } A$ is a subspace of \mathbb{R}^n .

The Null Space of a Matrix and the Kernel of a Matrix Transformation

In this section we explore the *null space* and see how the null space of a matrix is related to the matrix transformation defined by the matrix.

Let A be an $m \times n$ matrix. In Preview Activity 13.1 we defined the null space of a matrix A (see Definition 13.1) as the set of solutions to the matrix equation $A\mathbf{x} = \mathbf{0}$. Note that the null space of an $m \times n$ matrix is a subset of \mathbb{R}^n . We saw that the null space of A is closed under addition and scalar multiplication – that is, if \mathbf{u} and \mathbf{v} are in $\text{Nul } A$ and a and b are any scalars, then $\mathbf{u} + \mathbf{v}$ and $a\mathbf{u}$ are also in $\text{Nul } A$. Since the zero vector is always in $\text{Nul } A$, we can conclude that the null space of A is a subspace of \mathbb{R}^n .

There is a connection between the null space of a matrix and the matrix transformation it defines. Recall that any $m \times n$ matrix A defines a matrix transformation T from \mathbb{R}^n to \mathbb{R}^m by $T(\mathbf{x}) = A\mathbf{x}$. The null space of A is then the collection of vectors \mathbf{x} in \mathbb{R}^n so that $T(\mathbf{x}) = \mathbf{0}$. So if T is a matrix transformation from \mathbb{R}^n to \mathbb{R}^m , then the set

$$\{\mathbf{x} \in \mathbb{R}^n : T(\mathbf{x}) = \mathbf{0}\}$$

is a subspace of \mathbb{R}^n equal to the null space of A . This set is given a special name.

Definition 13.2. Let $T : \mathbb{R}^n \rightarrow \mathbb{R}^m$ be a matrix transformation. The **kernel** of T is the set

$$\text{Ker}(T) = \{\mathbf{x} \in \mathbb{R}^n : T(\mathbf{x}) = \mathbf{0}\}.$$



Activity 13.1. If T is a matrix transformation defined by a matrix A , then there is a convenient way to determine if T is one-to-one.

- (a) Let T be the matrix transformation defined by $T(\mathbf{x}) = A\mathbf{x}$, where

$$A = \begin{bmatrix} 1 & 2 & -1 \\ 0 & 1 & 4 \end{bmatrix}.$$

Find all of the vectors in $\text{Nul } A$. If $\text{Nul } A$ contains more than one vector, can T be one-to-one? Why?

- (b) Let T be the matrix transformation defined by $T(\mathbf{x}) = A\mathbf{x}$, where

$$A = \begin{bmatrix} 1 & 0 \\ 2 & 1 \\ -1 & 4 \end{bmatrix}.$$

Find all of the vectors in $\text{Nul } A$. Is T one-to-one? Why?

- (c) To find the vectors in the null space of a matrix A we solve the system $A\mathbf{x} = \mathbf{0}$. Since a homogeneous system is always consistent, there are two possibilities for $\text{Nul } A$: either $\text{Nul } A = \{\mathbf{0}\}$ or $\text{Nul } A$ contains infinitely many vectors.
- Under what conditions on A is $\text{Nul } A = \{\mathbf{0}\}$? What does that mean about T being one-to-one or not? Explain.
 - Under what conditions is $\text{Nul } A$ infinite? What does that mean about T being one-to-one or not? Explain.
 - Is it possible for $\text{Nul } A$ to be the whole space \mathbb{R}^n ? If so, give an example. If not, explain why not.

Recall that for a function to be one-to-one, each output must come from exactly one input. Since a matrix transformation T defined by $T(\mathbf{x}) = A\mathbf{x}$ always maps the zero vector to the zero vector, for T to be one-to-one it must be the case that the zero vector is the only vector that T maps to the zero vector. This means that the null space of A must be $\{\mathbf{0}\}$. Activity 13.1 demonstrates that if the matrix A that defines the transformation T has a pivot in every column, then $T(\mathbf{x}) = \mathbf{b}$ will have exactly one solution for each \mathbf{b} in the range of T . So a trivial null space is enough to characterize a one-to-one matrix transformation.

Theorem 13.3. *A matrix transformation T from \mathbb{R}^n to \mathbb{R}^m defined by $T(\mathbf{x}) = A\mathbf{x}$ is one-to-one if and only if*

$$\text{Nul } A = \text{Ker}(T) = \{\mathbf{0}\}.$$

The Column Space of a Matrix and the Range of a Matrix Transformation

Given an $m \times n$ matrix A , we have seen that the matrix-vector product $A\mathbf{x}$ is a linear combination of the columns of A with weights from \mathbf{x} . It follows that the equation $A\mathbf{x} = \mathbf{b}$ has a solution if and only if \mathbf{b} is a linear combination of the columns of A . So the span of the columns of A tells us for which vectors the equation $A\mathbf{x} = \mathbf{b}$ is consistent. We give the span of the columns of a matrix A a special name.



Definition 13.4. The **column space** of an $m \times n$ matrix A is the span of the columns of A .

We denote the column space of A as $\text{Col } A$. Given that $A\mathbf{x}$ is a linear combination of the columns of A , we can also write the column space of an $m \times n$ matrix A as

$$\text{Col } A = \{A\mathbf{x} : \mathbf{x} \text{ is in } \mathbb{R}^n\}.$$

For the matrix transformation T defined by $T(\mathbf{x}) = A\mathbf{x}$, the set of all vectors of the form $A\mathbf{x}$ is also the range of the transformation T . So for a matrix transformation T with matrix A we have $\text{Range}(T) = \text{Col } A$.

Activity 13.2. As a span of a set of vectors, we know that $\text{Col } A$ is a subspace of \mathbb{R}^k for an appropriate value of k .

- (a) Let $M = \begin{bmatrix} 1 & 1 & 1 & 0 & 2 \\ 0 & 1 & 0 & 1 & 1 \\ 1 & 0 & 1 & 1 & -1 \\ 0 & 1 & 0 & 1 & 1 \end{bmatrix}$. The space $\text{Col } M$ is a subspace of \mathbb{R}^k for some positive integer k . What is k in this case?
- (b) If A is an $m \times n$ matrix, then $\text{Col } A$ is a subspace of \mathbb{R}^k for some positive integer k . What is k in this case?
- (c) Recall that a matrix transformation T given by $T(\mathbf{x}) = A\mathbf{x}$ where A is an $m \times n$ matrix is onto if for every \mathbf{b} in \mathbb{R}^m there exists a \mathbf{x} in \mathbb{R}^n for which $T(\mathbf{x}) = \mathbf{b}$. How does T being onto relate to the $\text{Col } A$?

As you saw in Activity 13.2, a matrix transformation T defined by $T(\mathbf{x}) = A\mathbf{x}$ is onto if the column space of A , which consists of the image vectors under the transformation T , is equal to \mathbb{R}^m . In other words, we want the $\text{Range}(T)$ to equal \mathbb{R}^m .

Theorem 13.5. A matrix transformation T from \mathbb{R}^n to \mathbb{R}^m defined by $T(\mathbf{x}) = A\mathbf{x}$ is onto if and only if

$$\text{Col } A = \text{Range}(T) = \mathbb{R}^m.$$

The Row Space of a Matrix

As you might expect, if there is a column space for a matrix then there is also a row space for a matrix. The row space is defined just as the column space as the span of the rows of a matrix.

Definition 13.6. The **row space** of an $m \times n$ matrix A is the span of the row of A .

There is really nothing new here, though. Since the rows of A are the columns of A^\top , it follows that $\text{Row } A = \text{Col } A^\top$. So if we want to learn anything about the row space of A , we can just translate all of our questions to the column space of A^\top .



Bases for $\text{Nul } A$ and $\text{Col } A$

When confronted with a subspace of \mathbb{R}^n , we will usually want to find a minimal spanning set – a smallest spanning set – for the space. Recall that a minimal spanning set is also called a basis for the space. So a basis for a space must span that space, and to be a minimal spanning set we have seen that a basis must also be linearly independent. So to prove that a set is a basis for a subspace of \mathbb{R}^n we need to demonstrate two things: that the set is linearly independent, and that the set spans the subspace.

Activity 13.3. In this activity we see how to find a basis for $\text{Col } A$ and $\text{Nul } A$ for a specific matrix A . Let

$$A = \begin{bmatrix} 1 & 1 & 1 & 0 & 2 \\ 0 & 1 & 0 & 1 & 1 \\ 1 & 0 & 1 & 1 & -1 \\ 0 & 1 & 0 & 1 & 1 \end{bmatrix}.$$

Assume that the reduced row echelon form of A is

$$R = \begin{bmatrix} 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 2 \\ 0 & 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

- (a) First we examine $\text{Col } A$. Recall that to find a minimal spanning set of a set of vectors $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k\}$ in \mathbb{R}^n we just select the pivot columns of the matrix $[\mathbf{v}_1 \ \mathbf{v}_2 \ \dots \ \mathbf{v}_k]$.
 - i. Find a basis for $\text{Col } A$.
 - ii. Does $\text{Col } A$ equal $\text{Col } R$? Explain.
- (b) Now we look at $\text{Nul } A$.
 - i. Write the general solution to the homogeneous system $A\mathbf{x} = \mathbf{0}$ in vector form.
 - ii. Find a spanning set for $\text{Nul } A$.
 - iii. Find a basis for $\text{Nul } A$. Explain how you know you have a basis.

You should have noticed that Activity 13.3 (a) provides a process for finding a basis for $\text{Col } A$ – the pivot columns of A form a basis for $\text{Col } A$. Similarly, Activity 13.3 (b) shows us that we can find a basis for $\text{Nul } A$ by writing the general solution to the homogeneous equation $A\mathbf{x} = \mathbf{0}$ as a linear combination of vectors whose weights are the variables corresponding to the non-pivot columns of A – and these vectors form a basis for $\text{Nul } A$. As we will argue next, these process always give us bases for $\text{Col } A$ and $\text{Nul } A$.

Let A be an $m \times n$ matrix, and let R be the reduced row echelon form of A . Suppose R has k non-pivot columns and $n - k$ pivot columns. We can rearrange the columns so that the non-pivot columns of R are the last k columns (this just amounts to relabeling the unknowns in the system).

Basis for $\text{Nul } A$. Here we argue that the method described following Activity 13.3 to find a spanning set for the null space always yields a basis for the null space.



First note that $\text{Nul } R = \text{Nul } A$, since the system $Ax = \mathbf{0}$ has the same solution set as $Rx = \mathbf{0}$. So it is enough to find a basis for $\text{Nul } R$. If every column of R is a pivot column, then $Rx = \mathbf{0}$ has only the trivial solution and the null space of R is $\{\mathbf{0}\}$. Let us now consider the case where R contains non-pivot columns. If we let $\mathbf{x} = [x_1 \ x_2 \ \dots \ x_n]^T$, and if $Rx = \mathbf{0}$ then we can write x_1, x_2, \dots, x_{n-k} in terms of $x_{n-k+1}, x_{n-k+2}, \dots$, and x_n . From these equations we can write \mathbf{x} as a linear combination of some vectors $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k$ with $x_{n-k+1}, x_{n-k+2}, \dots, x_n$ as weights. By construction, each of the vectors $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k$ has a component that is 1 with the corresponding component as 0 in all the other \mathbf{v}_i . Therefore, the vectors $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k$ are linearly independent and span $\text{Nul } R$ (and $\text{Nul } A$). In other words, the method we have developed to find the general solution to $Ax = \mathbf{0}$ always produces a basis for $\text{Nul } A$.

Basis for Col A . Here we explain why the pivot columns of A form a basis for $\text{Col } A$. Recall that the product Ax expresses a linear combination of the columns of A with weights from \mathbf{x} , and every such linear combination is matched with a product Rx giving a linear combination of the columns of R using the same weights. So if a set of columns of R is linearly independent (or dependent), then the set of corresponding columns in A is linearly independent (or dependent) and vice versa. Since each pivot column of R is a vector with 1 in one entry (a different entry for different pivot columns) and zeros elsewhere, the pivot columns of R are clearly linearly independent. It follows that the pivot columns of A are linearly independent. All that remains is to explain why the pivot columns of A span $\text{Col } A$.

Let $\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_n$ be the columns of R so that $R = [\mathbf{r}_1 \ \mathbf{r}_2 \ \dots \ \mathbf{r}_n]$, and let $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n$ be the columns of A so that $A = [\mathbf{a}_1 \ \mathbf{a}_2 \ \dots \ \mathbf{a}_n]$. Suppose \mathbf{a}_i is a non-pivot column for A and \mathbf{r}_i the corresponding non-pivot column in R . Each pivot column is composed of a single 1 with the rest of its entries 0. Also, if a non-pivot column contains a nonzero entry, then there is a corresponding pivot column that contains a 1 in the corresponding position. So \mathbf{r}_i is a linear combination of $\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_{n-k}$ – the pivot columns of R . Thus,

$$\mathbf{r}_i = c_1 \mathbf{r}_1 + c_2 \mathbf{r}_2 + \dots + c_{n-k} \mathbf{r}_{n-k}$$

for some scalars c_1, c_2, \dots, c_{n-k} . Let $\mathbf{x} = [c_1 \ c_2 \ \dots \ c_{n-k} \ 0 \ \dots \ 0 \ -1 \ 0 \ \dots \ 0]^T$, where the -1 is in position i . Then $Rx = \mathbf{0}$ and so $Ax = \mathbf{0}$. Thus,

$$\mathbf{a}_i = c_1 \mathbf{a}_1 + c_2 \mathbf{a}_2 + \dots + c_{n-k} \mathbf{a}_{n-k}$$

and \mathbf{a}_i is a linear combination of the pivot columns of A . So every non-pivot column of A is in the span of A and we conclude that the pivot columns of A form a basis for $\text{Col } A$.

IMPORTANT POINT: It is the pivot columns of A that form a basis for $\text{Col } A$, not the pivot columns of the reduced row echelon form R of A . In general, $\text{Col } R \neq \text{Col } A$.

We can incorporate the ideas of this section to expand the Invertible Matrix Theorem.

Theorem 13.7 (The Invertible Matrix Theorem). *Let A be an $n \times n$ matrix. The following statements are equivalent.*

- (1) *The matrix A is an invertible matrix.*
- (2) *The matrix equation $Ax = \mathbf{0}$ has only the trivial solution.*
- (3) *The matrix A has n pivot columns.*



- (4) Every row of A contains a pivot.
- (5) The columns of A span \mathbb{R}^n .
- (6) The matrix A is row equivalent to the identity matrix I_n .
- (7) The columns of A are linearly independent.
- (8) The columns of A form a basis for \mathbb{R}^n .
- (9) The matrix transformation T from \mathbb{R}^n to \mathbb{R}^n defined by $T(\mathbf{x}) = A\mathbf{x}$ is one-to-one.
- (10) The matrix equation $A\mathbf{x} = \mathbf{b}$ has exactly one solution for each vector \mathbf{b} in \mathbb{R}^n .
- (11) The matrix transformation T from \mathbb{R}^n to \mathbb{R}^n defined by $T(\mathbf{x}) = A\mathbf{x}$ is onto.
- (12) There is an $n \times n$ matrix C so that $AC = I_n$.
- (13) There is an $n \times n$ matrix D so that $DA = I_n$.
- (14) The scalar 0 is not an eigenvalue of A .
- (15) The matrix A^\top is invertible.
- (16) $\text{Nul } A = \{\mathbf{0}\}$.
- (17) $\text{Col } A = \mathbb{R}^n$.

Examples

What follows are worked examples that use the concepts from this section.

Example 13.8.

(a) Let $A = \begin{bmatrix} 1 & 0 & -2 & 3 \\ -2 & -4 & 0 & -14 \\ 1 & 3 & 1 & 9 \end{bmatrix}$.

- i. Find a basis for $\text{Col } A$.
- ii. Describe $\text{Col } A$ geometrically (e.g., as a line, a plane, a union of lines, etc.) in the appropriate larger space.

(b) Let $B = \begin{bmatrix} 0 & -2 & 1 \\ -1 & 0 & 1 \\ 6 & -10 & -1 \\ 1 & -4 & 1 \end{bmatrix}$.

- i. Find a basis for $\text{Nul } B$.
- ii. Describe $\text{Nul } B$ geometrically (e.g., as a line, a plane, a union of lines, etc.) in the appropriate larger space.

Example Solution.



(a) We use $A = \begin{bmatrix} 1 & 0 & -2 & 3 \\ -2 & -4 & 0 & -14 \\ 1 & 3 & 1 & 9 \end{bmatrix}$.

i. Technology shows that the reduced row echelon form of A is

$$\begin{bmatrix} 1 & 0 & -2 & 3 \\ 0 & 1 & 1 & 2 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

The first two columns of A are the pivot columns of A . Since the pivot columns of A form a basis for $\text{Col } A$, a basis for $\text{Col } A$ is

$$\left\{ \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ -4 \\ 3 \end{bmatrix} \right\}.$$

ii. Let $\mathbf{v}_1 = \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}$ and $\mathbf{v}_2 = \begin{bmatrix} 0 \\ -4 \\ 3 \end{bmatrix}$. Since neither \mathbf{v}_1 nor \mathbf{v}_2 is a scalar multiple of the other, we see that $\text{Col } A$ is the span of two linearly independent vectors in \mathbb{R}^3 . Thus, we conclude that $\text{Col } A$ is the plane in \mathbb{R}^3 through the origin and the points $(1, -2, 1)$ and $(0, -4, 3)$.

(b) We use $B = \begin{bmatrix} 0 & -2 & 1 \\ -1 & 0 & 1 \\ 6 & -10 & -1 \\ 1 & -4 & 1 \end{bmatrix}$.

i. Technology shows that the reduced row echelon form of B is

$$\begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & -\frac{1}{2} \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

To find a basis for $\text{Nul } B$, we must solve the homogeneous equation $B\mathbf{x} = \mathbf{0}$. If

$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$ and $B\mathbf{x} = \mathbf{0}$, the reduced row echelon form of B shows that x_3 is free,



$x_2 = \frac{1}{2}x_3$, and $x_1 = x_3$. So

$$\begin{aligned}\mathbf{x} &= \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \\ &= \begin{bmatrix} x_3 \\ \frac{1}{2}x_3 \\ x_3 \end{bmatrix} \\ &= x_3 \begin{bmatrix} 1 \\ \frac{1}{2} \\ 1 \end{bmatrix}.\end{aligned}$$

Thus, a basis for $\text{Nul } B$ is

$$\left\{ \begin{bmatrix} 2 \\ 1 \\ 2 \end{bmatrix} \right\}.$$

ii. Since $\text{Nul } B = \text{Span} \left\{ \begin{bmatrix} 2 \\ 1 \\ 2 \end{bmatrix} \right\}$ is the span of one nonzero vector in \mathbb{R}^3 , we conclude that $\text{Nul } B$ is the line in \mathbb{R}^3 through the origin and the point $(2, 1, 2)$.

Example 13.9. Let $A = \begin{bmatrix} 1 & 3 \\ -1 & 2 \\ 0 & -2 \\ 5 & 6 \end{bmatrix}$, and let T be the matrix transformation defined by $T(\mathbf{x}) = A\mathbf{x}$.

- (a) What are the domain and codomain of T ? Why?
- (b) Find all vectors \mathbf{x} such that $T(\mathbf{x}) = \mathbf{0}$. How is this set of vectors related to $\text{Nul } A$? Explain.
- (c) Is T one-to-one? Explain.
- (d) Is T onto? If yes, explain why. If no, find a basis for the range of T .

Example Solution.

- (a) Recall that $A\mathbf{x}$ is a linear combination of the columns of A with weights from \mathbf{x} . So $A\mathbf{x}$ is defined only when the number of components of \mathbf{x} is equal to the number of columns of A . This explains why the domain of T is \mathbb{R}^2 . Also, since each output of T is a linear combination of the columns of A , the codomain of T is \mathbb{R}^4 .
- (b) The set of vectors \mathbf{x} such that $\mathbf{0} = T(\mathbf{x}) = A\mathbf{x}$ is the same as $\text{Nul } A$. The reduced row echelon form of A is

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

Since both columns of A are pivot columns, the columns of A are linearly independent. This implies that $A\mathbf{x} = \mathbf{0}$ has only the trivial solution. Therefore, the only vector \mathbf{x} such that $T(\mathbf{x}) = \mathbf{0}$ is the zero vector in \mathbb{R}^2 .

- (c) The previous part shows that $\text{Ker}(T) = \{\mathbf{0}\}$. This means that T is one-to-one by Theorem 13.3.
- (d) Recall that the range of T is the same as $\text{Col } A$. The reduced row echelon form of A has a row of zeros, so $A\mathbf{x} = \mathbf{b}$ is not consistent for every \mathbf{b} in \mathbb{R}^4 . We conclude that T is not onto. To find a basis for the range of T , we just need to find a basis for $\text{Col } A$. The pivot columns of A form such a basis, so a basis for the range of T is

$$\left\{ \begin{bmatrix} 1 \\ -1 \\ 0 \\ 5 \end{bmatrix}, \begin{bmatrix} 3 \\ 2 \\ -2 \\ 6 \end{bmatrix} \right\}.$$

Summary

- The null space of an $m \times n$ matrix A is the set of vectors \mathbf{x} in \mathbb{R}^n so that $A\mathbf{x} = \mathbf{0}$. In set notation

$$\text{Nul } A = \{\mathbf{x} : A\mathbf{x} = \mathbf{0}\}.$$

- The column space of a matrix A is the span of the columns of A .
- A subset W of \mathbb{R}^n is a subspace of \mathbb{R}^n if

- (1) $\mathbf{u} + \mathbf{v}$ is in W whenever \mathbf{u} and \mathbf{v} are in W (when this property is satisfied we say that W is *closed* under addition),
- (2) $a\mathbf{u}$ is in W whenever a is a scalar and \mathbf{u} is in W (when this property is satisfied we say that W is *closed* under multiplication by scalars),
- (3) $\mathbf{0}$ is in W .

- The null space of an $m \times n$ matrix is a subspace of \mathbb{R}^n while the column space of A is a subspace of \mathbb{R}^m .
- The span of any set of vectors in \mathbb{R}^n is a subspace of \mathbb{R}^n .
- The kernel of a matrix transformation $T : \mathbb{R}^n \rightarrow \mathbb{R}^m$ is the set

$$\text{Ker}(T) = \{\mathbf{x} \in \mathbb{R}^n : T(\mathbf{x}) = \mathbf{0}\}.$$

- The kernel of a matrix transformation T defined by $T(\mathbf{x}) = A\mathbf{x}$ is the same set as $\text{Nul } A$.
- The range of a matrix transformation $T : \mathbb{R}^n \rightarrow \mathbb{R}^m$ is the set

$$\text{Range}(T) = \{T(\mathbf{x}) : \mathbf{x} \in \mathbb{R}^n\}.$$

- The range of a matrix transformation T defined by $T(\mathbf{x}) = A\mathbf{x}$ is the same set as $\text{Col } A$.



- A basis for the null space of a matrix A can be found by writing the general solution to the homogeneous equation $A\mathbf{x} = \mathbf{0}$ as a linear combination of vectors whose weights are the variables corresponding to the non-pivot columns of A . The number of vectors in a basis for $\text{Nul } A$ is the number of non-pivot columns of A .
- The pivot columns of a matrix A form a basis for the column space of A .

Exercises

- (1) Find a basis for the null space and column space of the matrix

$$A = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 0 & 0 & 2 & -2 \\ 1 & 2 & 5 & 2 \end{bmatrix}.$$

Of which spaces are the null and column spaces of A subspaces?

- (2) If the column space of $\begin{bmatrix} 1 & 2 & -1 \\ 1 & 1 & 1 \\ 1 & 2 & c \end{bmatrix}$ has basis $\left\{ \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \\ 2 \end{bmatrix} \right\}$, what must c be?

- (3) If the null space of $\begin{bmatrix} 2 & 1 & a \\ 1 & 2 & b \end{bmatrix}$ has basis $\left\{ \begin{bmatrix} 2 \\ -1 \\ 1 \end{bmatrix} \right\}$, what must a and b be?

- (4) Find a matrix with at least four non-zero and distinct columns for which the column space has basis $\left\{ \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 2 \\ 3 \end{bmatrix} \right\}$.

- (5) Find a matrix with at least two rows whose null space has basis $\left\{ \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix} \right\}$.

- (6) Find a matrix whose column space has basis $\left\{ \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 2 \\ 3 \end{bmatrix} \right\}$ and whose null space has basis $\left\{ \begin{bmatrix} 2 \\ 1 \\ -1 \end{bmatrix} \right\}$.

- (7) If possible, find a 4×4 matrix whose column space does not equal \mathbb{R}^4 but whose null space equals $\{\mathbf{0}\}$. Explain your answer. If not possible, explain why not.

- (8) Label each of the following statements as True or False. Provide justification for your response.

- (a) **True/False** For a 3×4 matrix, the null space contains vectors other than the zero vector.



- (b) **True/False** For a 4×3 matrix, the null space contains vectors other than the zero vector.
- (c) **True/False** If $\text{Nul } A$ is not the zero subspace, then the transformation $\mathbf{x} \mapsto A\mathbf{x}$ is not one-to-one.
- (d) **True/False** If the transformation $\mathbf{x} \mapsto A\mathbf{x}$ is onto where A is an $m \times n$ matrix, then $\text{Col } A = \mathbb{R}^m$.
- (e) **True/False** For a 4×3 matrix A , $\text{Col } A$ cannot equal \mathbb{R}^4 .
- (f) **True/False** For a 3×4 matrix A , $\text{Col } A$ cannot equal \mathbb{R}^3 .
- (g) **True/False** The null space of the matrix $\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$ consists of the two vectors $[-1 \ 0 \ 1]^T$ and $[0 \ -1 \ 1]^T$.
- (h) **True/False** A basis for the null space of the matrix $\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$ consists of the two vectors $[-1 \ 0 \ 1]^T$ and $[0 \ -1 \ 1]^T$.
- (i) **True/False** There does not exist a matrix whose null space equals its column space.
- (j) **True/False** The column space of every 4×4 matrix is \mathbb{R}^4 and its null space is $\{\mathbf{0}\}$.

Project: Solving the Lights Out Game

The Lights Out game starts with a 5×5 grid on which some of the squares are lit (on) and some are not lit (off). We will call such a state a *configuration*. Pressing a square that is on turns it off and changes the state of all adjacent (vertically and horizontally) squares, and pressing a square that is off turns it on and changes the state of all adjacent (vertically and horizontally) squares. To model this situation, we consider the number system $\mathbb{Z}_2 = \{0, 1\}$ consisting only of 0 and 1, where 0 represents the off state and 1 the on state. We can also think of 1 as the act of pressing a square and 0 as the act of not pressing – that is,

- $0 + 0 = 0$ (not pressing an off square leaves it off),
- $0 + 1 = 1 + 0 = 1$ (pressing an off square turns it on or not pressing a lit square leaves it lit),
- $1 + 1 = 0$ (pressing a lit square turns it off).

The numbers 0 and 1 in \mathbb{Z}_2 will be the only numbers we use when playing the Lights Out game, so all of our matrix entries will be in \mathbb{Z}_2 and all of our calculations are done in \mathbb{Z}_2 .

There are two ways we can view a Lights Out game.

- We can view each configuration as a 5×5 matrix. In this situation, we label the entries in the grid as shown in Figure 13.1. Each entry in the grid will be assigned a 0 or 1 according to whether the light in that entry is off or on.



- For our purposes a better way to visualize a Lights Out configuration is as a 25×1 vector. The components in this vector correspond to the entries in the 5×5 grid with the correspondence given by the numbering demonstrated in Figure 13.1 (for the sake of space, this array is shown in a row instead of a column). Again, each component is assigned a 0 or 1 according to whether the light for that entry is off or on. In this view, each configuration is a vector with 25 components in \mathbb{Z}_2 .

1	2	3	4	5
6	7	8	9	10
11	12	13	14	15
16	17	18	19	20
21	22	23	24	25

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
---	---	---	---	---	---	---	---	---	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----

Figure 13.1: Two representations of the Lights Out game.

We will take the latter perspective and view the Lights Out game as if it is played on a 25×1 board with entries in \mathbb{Z}_2 . The space of all of these Lights Out configurations is denoted as \mathbb{Z}_2^{25} (similar to \mathbb{R}^{25} , but with entries in \mathbb{Z}_2 rather than \mathbb{R}). Since \mathbb{Z}_2 is a field, the space \mathbb{Z}_2^{25} is a vector space just as \mathbb{R}^{25} is. This is the environment in which we will play the Lights Out game.

If we think about the game as played on a 25×1 board, then pressing a square correlates to selecting one of the 25 components of a configuration vector. Each time we press a square, we make a move that changes the status of that square and all the squares vertically or horizontally adjacent to it from the 5×5 board. Recalling that adding 1 to a square has the effect of changing its status (from on to off or off to on), and each move that we make in the game can be represented as a 25×1 vector that is added to a configuration. For example, the move of pressing the first square is given by adding the vector

$$\mathbf{m}_1 = [1 \ 1 \ 0 \ 0 \ 0 \ 1 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0]^T$$

to a configuration vector and the move of pressing the second square is represented by adding the vector

$$\mathbf{m}_2 = [1 \ 1 \ 1 \ 0 \ 0 \ 0 \ 1 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0]^T.$$

Project Activity 13.1. Let \mathbf{m}_i be the move of pressing the i th square for i from 1 to 25.

- Find vector representations for \mathbf{m}_9 and \mathbf{m}_{22} .
- Let $M = [m_{ij}] = [\mathbf{m}_1 | \mathbf{m}_2 | \cdots | \mathbf{m}_{25}]$. Explain why $m_{ij} = m_{ji}$ for each i and j . In other words, explain why $M^T = M$. (Such a matrix is called a *symmetric* matrix.)

The goal of the Lights Out game is to begin with an initial configuration \mathbf{c} (a vector in \mathbb{Z}_2^{25}) and determine if we can apply a sequence of moves to obtain the configuration in which all the entries are 0 (or all the lights are off). The vector in \mathbb{Z}_2^{25} of all 0s is the zero vector in \mathbb{Z}_2^{25} and we will denote it as $\mathbf{0}$. Some basic algebra of vector addition in \mathbb{Z}_2 (or mod 2) will help us understand the strategy.



Start with a configuration \mathbf{c} . If we press the i th square, then we obtain the new configuration $\mathbf{c}_1 = \mathbf{m}_i + \mathbf{c}$ (where each move \mathbf{m}_i is also in \mathbb{Z}_2^{25}).

Project Activity 13.2.

- What happens if we press the i th square twice in a row? Explain in terms of the action and the game and verify using vector addition.
- Explain why applying move \mathbf{m}_i then move \mathbf{m}_j is the same as applying move \mathbf{m}_j , then \mathbf{m}_i .
- Explain how the answers to the previous two questions show that to play the game we only need to determine which buttons to press (and only once each) without worrying about the order in which the buttons are pressed.

What we have seen is that to play the game we are really looking for scalars x_1, x_2, \dots, x_{25} in \mathbb{Z}_2 (in other words, either 0 or 1) so that

$$x_1\mathbf{m}_1 + x_2\mathbf{m}_2 + \cdots + x_{25}\mathbf{m}_{25} + \mathbf{c} = \mathbf{0}. \quad (13.1)$$

Project Activity 13.3. Explain why (13.1) has the equivalent matrix equation

$$M\mathbf{x} = \mathbf{c}, \quad (13.2)$$

where $M = [\mathbf{m}_1 | \mathbf{m}_2 | \cdots | \mathbf{m}_{25}]$, or

$$M = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ M = & \vdots & & & & & & & & & & & & & & & & & & & & & & & & & & & & & . \end{bmatrix}$$

Explicitly identify the vector \mathbf{x} . Also, explain why \mathbf{c} is on the right side of this equation.



To solve a Lights Out game now, all we need do is determine a solution, if one exists, to the matrix equation (13.2).

Project Activity 13.4. For this activity you may use the fact that the reduced row echelon form of the matrix M (using algebra in \mathbb{Z}_2) is as shown below.

- Find a basis for the column space of M .
- Explain why not every Lights Out puzzle can be solved. That is, explain why there are some initial configurations of lights on and off for which it is not possible to turn out all the lights (without turning off the game). Relate this to the column space of M .

The reduced row echelon form of the matrix M (using algebra in \mathbb{Z}_2):

$$\left[\begin{array}{cccccccccccccccccccccc} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{array} \right] .$$

To find conditions under which a Lights Out game is not solvable, we will demonstrate that if A is an $n \times n$ matrix, then the scalar product of any vector in $\text{Nul } A^\top$ with any column of A is $\mathbf{0}$. Let $A = [a_{ij}]$ be an $n \times n$ matrix with columns $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n$. Represent the entries in the i th column as $\mathbf{a}_i = [a_{1i} \ a_{2i} \ \dots \ a_{ni}]^\top$ for each i between 1 and n . Note that \mathbf{a}_i is also the i th row of A^\top . Also, let $\mathbf{x} = [x_1 \ x_2 \ \dots \ x_n]^\top$ be a vector in $\text{Nul } A^\top$. Then $A^\top \mathbf{x} = \mathbf{0}$. Using the row-column method of multiplying a matrix by a vector, when we multiply the i th row of A^\top with \mathbf{x} we obtain

$$a_{1i}x_1 + a_{2i}x_2 + \dots + a_{ni}x_n = 0. \quad (13.3)$$



This equation is valid for each i between 1 and n . Recall that the sum in (13.3) is the scalar product of \mathbf{a}_i and \mathbf{x} and is denoted $\mathbf{a}_i \cdot \mathbf{x}$. That is,

$$\mathbf{a}_i \cdot \mathbf{x} = a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n.$$

The fact that \mathbf{x} is in $\text{Nul } A^\top$ means $\mathbf{a}_i \cdot \mathbf{x} = \mathbf{0}$ for every i between 1 and n . In other words, the scalar product of any vector in $\text{Nul } A^\top$ with any column of A is $\mathbf{0}$. (When the scalar product of two vectors is $\mathbf{0}$, we call the vectors *orthogonal* – a fancy word for “perpendicular”.) Since scalar products are linear, we can extend this result to the following.

Theorem 13.10. *Let A be an $n \times n$ matrix. If \mathbf{x} is any vector in $\text{Col } A$ and \mathbf{y} is any vector in $\text{Nul } A^\top$, then $\mathbf{x} \cdot \mathbf{y} = 0$.*

With Theorem 13.10 in mind we can return to our analysis of the Lights Out game, applying this result to the matrix M .

Project Activity 13.5.

- (a) Find a basis for the null space of M^\top . (Recall that $M^\top = M$, so you can use the reduced row echelon form of M (using algebra in \mathbb{Z}_2) given earlier.)
- (b) Use Theorem 13.10 to show that if $\mathbf{c} = [c_1 \ c_2 \ \dots \ c_{25}]^\top$ is an initial Lights Out configuration that is solvable, then \mathbf{c} must be orthogonal to each of the vectors in a basis for $\text{Nul } M^\top$. Then show that if \mathbf{c} is a solvable initial Lights Out configuration, \mathbf{c} must satisfy

$$\begin{aligned} c_2 + c_3 + c_4 + c_6 + c_8 + c_{10} + c_{11} + c_{12} + c_{14} + c_{15} + c_{16} + c_{18} \\ + c_{20} + c_{22} + c_{23} + c_{24} = 0 \end{aligned}$$

and

$$c_1 + c_3 + c_5 + c_6 + c_8 + c_{10} + c_{11} + c_{13} + c_{15} + c_{17} + c_{19} + c_{21} + c_{23} + c_{25} = 0.$$

Be very specific in your explanation.

Project Activity 13.6. Now that we know which Lights Out games can be solved, let \mathbf{c} be an initial configuration to a solvable Lights Out game. Explain how to find a solution to this game. Will the solution be unique? Explain.

Now that we have a strategy for solving the Lights Out game, use it to solve random puzzles at <https://www.geogebra.org/m/wcmctahp>, or create your own game to solve.

Section 14

Eigenspaces of a Matrix

Focus Questions

By the end of this section, you should be able to give precise and thorough answers to the questions listed below. You may want to keep these questions in mind to focus your thoughts as you complete the section.

- What is an eigenspace of a matrix?
- How do we find a basis for an eigenspace of a matrix?
- What is true about any set of eigenvectors for a matrix that correspond to different eigenvalues?

Application: Population Dynamics

The study of population dynamics – how and why people move from one place to another – is important to economists. The movement of people corresponds to the movement of money, and money makes the economy go. As an example, we might consider a simple model of population migration to and from the state of Michigan.

According to the Michigan Department of Technology, Management, and Budget,¹ from 2011 to 2012, approximately 0.05% of the U.S. population outside of Michigan moved to the state of Michigan, while approximately 2% of Michigan's population moved out of Michigan. A reasonable question to ask about this situation is, if these numbers don't change, what is the long-term distribution of the US population inside and outside of Michigan (under the assumption that the total US population doesn't change.). The answer to this question involves eigenvalues and eigenvectors of a matrix. More details can be found later in this section.

¹<http://michigan.gov/cgi/0,1607,7-158-54534-140915--,00.html>

Introduction

Preview Activity 14.1. Consider the matrix transformation T from \mathbb{R}^2 to \mathbb{R}^2 defined by $T(\mathbf{x}) = A\mathbf{x}$, where

$$A = \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix}.$$

We are interested in understanding what this matrix transformation does to vectors in \mathbb{R}^2 . The matrix A has eigenvalues $\lambda_1 = 2$ and $\lambda_2 = 4$ with corresponding eigenvectors $\mathbf{v}_1 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$ and $\mathbf{v}_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$.

- (1) Explain why \mathbf{v}_1 and \mathbf{v}_2 are linearly independent.
- (2) Explain why any vector \mathbf{b} in \mathbb{R}^2 can be written uniquely as a linear combination of \mathbf{v}_1 and \mathbf{v}_2 .
- (3) We now consider the action of the matrix transformation T on a linear combination of \mathbf{v}_1 and \mathbf{v}_2 . Explain why

$$T(c_1\mathbf{v}_1 + c_2\mathbf{v}_2) = 2c_1\mathbf{v}_1 + 4c_2\mathbf{v}_2. \quad (14.1)$$

Equation (14.1) illustrates that it would be convenient to view the action of T in the coordinate system where $\text{Span}\{\mathbf{v}_1\}$ serves as the x -axis and $\text{Span}\{\mathbf{v}_2\}$ as the y -axis. In this case, we can visualize that when we apply the transformation T to a vector $\mathbf{b} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2$ in \mathbb{R}^2 the result is an output vector scaled by a factor of 2 in the \mathbf{v}_1 direction and by a factor of 4 in the \mathbf{v}_2 direction. For example, consider the box with vertices at $(0, 0)$, \mathbf{v}_1 , \mathbf{v}_2 , and $\mathbf{v}_1 + \mathbf{v}_2$ as shown at left in Figure 14.1. The transformation T stretches this box by a factor of 2 in the \mathbf{v}_1 direction and a factor of 4 in the \mathbf{v}_2 direction as illustrated at right in Figure 14.1. In this situation, the eigenvalues and eigenvectors provide the most convenient perspective through which to visualize the action of the transformation T . Here, $\text{Span}\{\mathbf{v}_1\}$ and $\text{Span}\{\mathbf{v}_2\}$ are the eigenspaces of the matrix A .

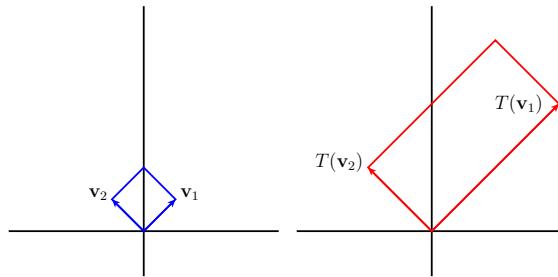


Figure 14.1: A box and a transformed box.

This geometric perspective illustrates how each the span of each eigenvalue of A tells us something important about A . In this section we explore the idea of eigenvalues and spaces defined by eigenvectors in more detail.

Eigenspaces of Matrix

Recall that the eigenvectors of an $n \times n$ matrix A satisfy the equation

$$A\mathbf{x} = \lambda\mathbf{x}$$

for some scalar λ . Equivalently, the eigenvectors of A with eigenvalue λ satisfy the equation

$$(A - \lambda I_n)\mathbf{x} = \mathbf{0}.$$

In other words, the eigenvectors for A with eigenvalue λ are the non-zero vectors in $\text{Nul } A - \lambda I_n$. Recall that the null space of an $n \times n$ matrix is a subspace of \mathbb{R}^n . In Preview Activity 14.1 we say how these subspaces provided a convenient coordinate system through which to view a matrix transformation. These special null spaces are called *eigenspaces*.

Definition 14.1. Let A be an $n \times n$ matrix with eigenvalue λ . The **eigenspace** for A corresponding to λ is the null space of $A - \lambda I_n$.

Activity 14.1. The matrix $A = \begin{bmatrix} 2 & 0 & 1 \\ 0 & 2 & -1 \\ 0 & 0 & 1 \end{bmatrix}$ has two distinct eigenvalues.

- (a) Find a basis for the eigenspace of A corresponding to the eigenvalue $\lambda_1 = 1$. In other words, find a basis for $\text{Nul } A - I_3$.
- (b) Find a basis for the eigenspace of A corresponding to the eigenvalue $\lambda_2 = 2$.
- (c) Is it true that if \mathbf{v}_1 and \mathbf{v}_2 are two distinct eigenvectors for A , that \mathbf{v}_1 and \mathbf{v}_2 are linearly independent? Explain.
- (d) Is it possible to have two linearly independent eigenvectors corresponding to the same eigenvalue?
- (e) Is it true that if \mathbf{v}_1 and \mathbf{v}_2 are two distinct eigenvectors corresponding to different eigenvalues for A , that \mathbf{v}_1 and \mathbf{v}_2 are linearly independent? Explain.

If we know an eigenvalue λ of an $n \times n$ matrix A , Activity 14.1 shows us how to find a basis for the corresponding eigenspace – just row reduce $A - \lambda I_n$ to find a basis for $\text{Nul } A - \lambda I_n$. To this point we have always been given eigenvalues for our matrices, and have not seen how to find these eigenvalues. That process will come a bit later. For now, we just want to become more familiar with eigenvalues and eigenvectors. The next activity should help connect eigenvalues to ideas we have discussed earlier.

Activity 14.2. Let A be an $n \times n$ matrix with eigenvalue λ .

- (a) How many solutions does the equation $(A - \lambda I_n)\mathbf{x} = \mathbf{0}$ have? Explain.
- (b) Can $A - \lambda I_n$ have a pivot in every column? Why or why not?
- (c) Can $A - \lambda I_n$ have a pivot in every row? Why or why not?
- (d) Can the columns of $A - \lambda I_n$ be linearly independent? Why or why not?



Linearly Independent Eigenvectors

An important question we will want to answer about a matrix is how many linearly independent eigenvectors the matrix has. Activity 14.1 shows that eigenvectors for the same eigenvalue may be linearly dependent or independent, but all of our examples so far seem to indicate that eigenvectors corresponding to different eigenvalues are linearly independent. This turns out to be universally true as our next theorem demonstrates. The next activity should help prepare us for the proof of this theorem.

Activity 14.3. Let λ_1 and λ_2 be distinct eigenvalues of a matrix A with corresponding eigenvectors \mathbf{v}_1 and \mathbf{v}_2 . The goal of this activity is to demonstrate that \mathbf{v}_1 and \mathbf{v}_2 are linearly independent. To prove that \mathbf{v}_1 and \mathbf{v}_2 are linearly independent, suppose that

$$x_1\mathbf{v}_1 + x_2\mathbf{v}_2 = \mathbf{0}. \quad (14.2)$$

- (a) Multiply both sides of equation (14.2) on the left by the matrix A and show that

$$x_1\lambda_1\mathbf{v}_1 + x_2\lambda_2\mathbf{v}_2 = \mathbf{0}. \quad (14.3)$$

- (b) Now multiply both sides of equation (14.2) by the scalar λ_1 and show that

$$x_1\lambda_1\mathbf{v}_1 + x_2\lambda_1\mathbf{v}_2 = \mathbf{0}. \quad (14.4)$$

- (c) Combine equations (14.3) and (14.4) to obtain the equation

$$x_2(\lambda_2 - \lambda_1)\mathbf{v}_2 = \mathbf{0}. \quad (14.5)$$

- (d) Explain how we can conclude that $x_2 = 0$. Why does it follow that $x_1 = 0$? What does this tell us about \mathbf{v}_1 and \mathbf{v}_2 ?

Activity 14.3 contains the basic elements of the proof of the next theorem.

Theorem 14.2. Let $\lambda_1, \lambda_2, \dots, \lambda_k$ be k distinct eigenvalues for a matrix A and for each i between 1 and k let \mathbf{v}_i be an eigenvector of A with eigenvalue λ_i . Then the vectors $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k$ are linearly independent.

Proof. Let A be a matrix with k distinct eigenvalues $\lambda_1, \lambda_2, \dots, \lambda_k$ and corresponding eigenvectors $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k$. To understand why $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k$ are linearly independent, we will argue by contradiction and suppose that the vectors $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k$ are linearly dependent. Note that \mathbf{v}_1 cannot be the zero vector (why?), so the set $S_1 = \{\mathbf{v}_1\}$ is linearly independent. If we include \mathbf{v}_2 into this set, the set $S_2 = \{\mathbf{v}_1, \mathbf{v}_2\}$ may be linearly independent or dependent. If S_2 is linearly independent, then the set $S_3 = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ may be linearly independent or dependent. We can continue adding additional vectors until we reach the set $S_k = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \dots, \mathbf{v}_k\}$ which we are assuming is linearly dependent. So there must be a smallest integer $m \geq 2$ such that the set S_m is linearly dependent while S_{m-1} is linearly independent. Since $S_m = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \dots, \mathbf{v}_m\}$ is



linearly dependent, there is a linear combination of $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_m$ with weights not all 0 that is the zero vector. Let c_1, c_2, \dots, c_m be such weights, not all zero, so that

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_{m-1}\mathbf{v}_{m-1} + c_m\mathbf{v}_m = \mathbf{0} \quad (14.6)$$

If we multiply both sides of (14.6) on the left by the matrix A we obtain

$$\begin{aligned} A(c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_m\mathbf{v}_m) &= A\mathbf{0} \\ c_1A\mathbf{v}_1 + c_2A\mathbf{v}_2 + \cdots + c_mA\mathbf{v}_m &= \mathbf{0} \\ c_1\lambda_1\mathbf{v}_1 + c_2\lambda_2\mathbf{v}_2 + \cdots + c_m\lambda_m\mathbf{v}_m &= \mathbf{0}. \end{aligned} \quad (14.7)$$

If we multiply both sides of (14.6) by λ_m we obtain the equation

$$c_1\lambda_m\mathbf{v}_1 + c_2\lambda_m\mathbf{v}_2 + \cdots + c_m\lambda_m\mathbf{v}_m = \mathbf{0}. \quad (14.8)$$

Subtracting corresponding sides of equation (14.8) from (14.7) gives us

$$c_1(\lambda_1 - \lambda_m)\mathbf{v}_1 + c_2(\lambda_2 - \lambda_m)\mathbf{v}_2 + \cdots + c_{m-1}(\lambda_{m-1} - \lambda_m)\mathbf{v}_{m-1} = \mathbf{0}. \quad (14.9)$$

Recall that S_{m-1} is a linearly independent set, so the only way a linear combination of vectors in S_{m-1} can be $\mathbf{0}$ is if all of the weights are 0. Therefore, we must have

$$c_1(\lambda_1 - \lambda_m) = 0, \quad c_2(\lambda_2 - \lambda_m) = 0, \quad \dots, \quad c_{m-1}(\lambda_{m-1} - \lambda_m) = 0.$$

Since the eigenvalues are all distinct, this can only happen if

$$c_1 = c_2 = \cdots = c_{m-1} = 0.$$

But equation (14.6) then implies that $c_m = 0$ and so all of the weights c_1, c_2, \dots, c_m are 0. However, when we assumed that the eigenvectors $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k$ were linearly dependent, this led to having at least one of the weights c_1, c_2, \dots, c_m be nonzero. This cannot happen, so our assumption that the eigenvectors $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k$ were linearly dependent must be false and we conclude that the eigenvectors $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k$ are linearly independent. ■

Examples

What follows are worked examples that use the concepts from this section.

Example 14.3. Let $A = \begin{bmatrix} 4 & -3 & -3 \\ -3 & 4 & 3 \\ 3 & -3 & -2 \end{bmatrix}$ and let T be the matrix transformation defined by $T(\mathbf{x}) = A\mathbf{x}$.

- (a) Show that 4 is an eigenvalue for A and find a basis for the corresponding eigenspace of A .
- (b) Geometrically describe the eigenspace of A corresponding to the eigenvalue 4. Explain what the transformation T does to this eigenspace.



- (c) Show that 1 is an eigenvalue for A and find a basis for the corresponding eigenspace of A .
 (d) Geometrically describe the eigenspace of A corresponding to the eigenvalue 1. Explain what the transformation T does to this eigenspace.

Example Solution.

- (a) Recall that λ is an eigenvalue of A if $A - \lambda I_3$ is not invertible. To show that 4 is an eigenvalue for A we row reduce the matrix

$$A - (4)I_3 = \begin{bmatrix} 0 & -3 & -3 \\ 3 & 0 & -3 \\ 3 & -3 & -6 \end{bmatrix}$$

to $\begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix}$. Since the third column of $A - 4I_3$ is not a pivot column, the matrix $A - 4I_3$ is not invertible. We conclude that 4 is an eigenvalue of A .

The eigenspace of A for the eigenvalue 4 is $\text{Nul}(A - 4I_3)$. The reduced row echelon form of $A - 4I_3$ shows that if $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$ and $(A - 4I_3)\mathbf{x} = \mathbf{0}$, then x_3 is free, $x_2 = -x_3$, and $x_1 = x_3$. Thus,

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_3 \\ -x_3 \\ x_3 \end{bmatrix} = x_3 \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}.$$

Therefore, $\left\{ \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} \right\}$ is a basis for the eigenspace of A corresponding to the eigenvalue 4.

- (b) Since the eigenspace of A corresponding to the eigenvalue 4 is the span of a single nonzero vector $\mathbf{v} = \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}$, this eigenspace is the line in \mathbb{R}^3 through the origin and the point $(1, -1, 1)$. Any vector in this eigenspace has the form $c\mathbf{v}$ for some scalar c . Notice that

$$T(c\mathbf{v}) = A(c\mathbf{v}) = c(A\mathbf{v}) = 4c\mathbf{v},$$

so T expands any vector in this eigenspace by a factor of 4.

- (c) To show that 1 is an eigenvalue for A we row reduce the matrix

$$A - (1)I_3 = \begin{bmatrix} 3 & -3 & -3 \\ -3 & 3 & 3 \\ 3 & -3 & -3 \end{bmatrix}$$

to $\begin{bmatrix} 1 & -1 & -1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$. Since the third column of $A - I_3$ is not a pivot column, the matrix $A - I_3$ is not invertible. We conclude that 1 is an eigenvalue of A .

The eigenspace of A for the eigenvalue 1 is $\text{Nul}(A - I_3)$. The reduced row echelon form of $A - I_3$ shows that if $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$ and $(A - I_3)\mathbf{x} = \mathbf{0}$, then x_2 and x_3 are free, and $x_1 = x_2 + x_3$. Thus,

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_2 + x_3 \\ x_2 \\ x_3 \end{bmatrix} = x_2 \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}.$$

Therefore, $\left\{ \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \right\}$ is a basis for the eigenspace of A corresponding to the eigenvalue 1.

- (d) Since the eigenspace of A corresponding to the eigenvalue 1 is the span of two linearly independent vectors $\mathbf{v}_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$ and $\mathbf{v}_2 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$, this eigenspace is the plane in \mathbb{R}^3 through the origin and the points $(1, 1, 0)$ and $(1, 0, 1)$. Any vector in this eigenspace has the form $a\mathbf{v}_1 + b\mathbf{v}_2$ for some scalars a and b . Notice that

$$T(a\mathbf{v}_1 + b\mathbf{v}_2) = A(a\mathbf{v}_1 + b\mathbf{v}_2) = aA\mathbf{v}_1 + bA\mathbf{v}_2 = a\mathbf{v}_1 + b\mathbf{v}_2,$$

so T fixes every vector in this plane.

Example 14.4.

- (a) Let $A = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}$. Note that the vector $\mathbf{v} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ satisfies $A\mathbf{v} = 3\mathbf{v}$.
- i. Show that \mathbf{v} is an eigenvector of A^2 . What is the corresponding eigenvalue?
 - ii. Show that \mathbf{v} is an eigenvector of A^3 . What is the corresponding eigenvalue?
 - iii. Show that \mathbf{v} is an eigenvector of A^4 . What is the corresponding eigenvalue?
 - iv. If k is a positive integer, do you expect that \mathbf{v} is an eigenvector of A^k ? If so, what do you think is the corresponding eigenvalue?
- (b) The result of part (a) is true in general. Let M be an $n \times n$ matrix with eigenvalue λ and corresponding eigenvector \mathbf{x} .
- i. Show that λ^2 is an eigenvalue of M^2 with eigenvector \mathbf{x} .
 - ii. Show that λ^3 is an eigenvalue of M^3 with eigenvector \mathbf{x} .
 - iii. Suppose that λ^k is an eigenvalue of M^k with eigenvector \mathbf{x} for some integer $k \geq 1$. Show then that λ^{k+1} is an eigenvalue of M^{k+1} with eigenvector \mathbf{x} . This argument shows that λ^k is an eigenvalue of M^k with eigenvector \mathbf{x} for any positive integer k .
- (c) We now investigate the eigenvalues of a special type of matrix.



- i. Let $B = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$. Show that $B^3 = 0$. (A square matrix M is *nilpotent*) if $M^k = 0$ for some positive integer k , so B is an example of a nilpotent matrix.) What are the eigenvalues of B ? Explain.
- ii. Show that the only eigenvalue of a nilpotent matrix is 0.

Example Solution.

(a) We use the fact that \mathbf{v} is an eigenvector of the matrix A with eigenvalue 3.

- i. We have that

$$A^2\mathbf{v} = A(A\mathbf{v}) = A(3\mathbf{v}) = 3(A\mathbf{v}) = 3(3\mathbf{v}) = 9\mathbf{v}.$$

So \mathbf{v} is an eigenvector of A^2 with eigenvalue $9 = 3^2$.

- ii. We have that

$$A^3\mathbf{v} = A(A^2\mathbf{v}) = A(9\mathbf{v}) = 9(A\mathbf{v}) = 9(3\mathbf{v}) = 27\mathbf{v}.$$

So \mathbf{v} is an eigenvector of A^3 with eigenvalue $27 = 3^3$.

- iii. We have that

$$A^4\mathbf{v} = A(A^3\mathbf{v}) = A(27\mathbf{v}) = 27(A\mathbf{v}) = 27(3\mathbf{v}) = 81\mathbf{v}.$$

So \mathbf{v} is an eigenvector of A^4 with eigenvalue $81 = 3^4$.

- iv. The results of the previous parts of this example indicate that $A^k\mathbf{v} = 3^k\mathbf{v}$, or that \mathbf{v} is an eigenvector of A^k with corresponding eigenvalue 3^k .

(b) Let M be an $n \times n$ matrix with eigenvalue λ and corresponding eigenvector \mathbf{x} .

- i. We have that

$$M^2\mathbf{x} = M(M\mathbf{x}) = M(\lambda\mathbf{x}) = \lambda(M\mathbf{x}) = \lambda(\lambda\mathbf{x}) = \lambda^2\mathbf{x}.$$

So \mathbf{x} is an eigenvector of M^2 with eigenvalue λ^2 .

- ii. We have that

$$M^3\mathbf{x} = M(M^2\mathbf{x}) = M(\lambda^2\mathbf{x}) = \lambda^2(M\mathbf{x}) = \lambda^2(\lambda\mathbf{x}) = \lambda^3\mathbf{x}.$$

So \mathbf{x} is an eigenvector of M^3 with eigenvalue λ^3 .

- iii. Assume that $M^k\mathbf{x} = \lambda^k\mathbf{x}$. Then

$$M^{k+1}\mathbf{x} = M(M^k\mathbf{x}) = M(\lambda^k\mathbf{x}) = \lambda^k(M\mathbf{x}) = 2\lambda^k(\lambda\mathbf{x}) = \lambda^{k+1}\mathbf{x}.$$

So \mathbf{x} is an eigenvector of M^{k+1} with eigenvalue λ^{k+1} .

(c) Now we investigate a special type of matrix.



- i. Straightforward calculations show that $B^3 = 0$. Since B is an upper triangular matrix, the eigenvalues of B are the entries on the diagonal. That is, the only eigenvalue of B is 0.
- ii. Assume that M is a nilpotent matrix. Suppose that λ is an eigenvalue of M with corresponding eigenvector v . Since M is a nilpotent matrix, there is a positive integer k such that $M^k = 0$. But λ^k is an eigenvalue of M^k with eigenvector v . The only eigenvalue of the zero matrix is 0, so $\lambda^k = 0$. This implies that $\lambda = 0$. We conclude that the only eigenvalue of a nilpotent matrix is 0.

Summary

- An eigenspace of an $n \times n$ matrix A corresponding to an eigenvalue λ of A is the null space of $A - \lambda I_n$.
- To find a basis for an eigenspace of a matrix A corresponding to an eigenvalue λ , we row reduce $A - \lambda I_n$ and find a basis for $\text{Nul } A - \lambda I_n$.
- Eigenvectors corresponding to different eigenvalues are always linearly independent.

Exercises

- (1) For each of the following, find a basis for the eigenspace of the indicated matrix corresponding to the given eigenvalue.

(a) $\begin{bmatrix} 10 & 7 \\ -14 & -11 \end{bmatrix}$ with eigenvalue 3

(b) $\begin{bmatrix} 11 & 18 \\ -3 & -4 \end{bmatrix}$ with eigenvalue 2

(c) $\begin{bmatrix} 2 & 1 \\ -1 & 0 \end{bmatrix}$ with eigenvalue 1

(d) $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 2 \\ 1 & 0 & 2 \end{bmatrix}$ with eigenvalue 2

(e) $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 2 \\ 1 & 0 & 2 \end{bmatrix}$ with eigenvalue 1

(f) $\begin{bmatrix} 2 & 2 & 4 \\ 1 & 1 & 2 \\ 3 & 3 & 6 \end{bmatrix}$ with eigenvalue 0

- (2) Suppose A is an invertible matrix.

- (a) Use the definition of an eigenvalue and an eigenvector to algebraically explain why if λ is an eigenvalue of A , then λ^{-1} is an eigenvalue of A^{-1} .



- (b) To provide an alternative explanation to the result in the previous part, let \mathbf{v} be an eigenvector of A corresponding to λ . Consider the matrix transformation T_A corresponding to A and $T_{A^{-1}}$ corresponding to A^{-1} . Considering what happens to \mathbf{v} if T_A and then $T_{A^{-1}}$ are applied, describe why this justifies \mathbf{v} is also an eigenvector of A^{-1} .

(3) If $A = \begin{bmatrix} 0 & 1 \\ a & b \end{bmatrix}$ has two eigenvalues 4 and 6, what are the values of a and b ?

(4)

- (a) What are the eigenvalues of the identity matrix I_2 ? Describe each eigenspace.
- (b) Now let $n > 2$ be a positive integer. What are the eigenvalues of the identity matrix I_n ? Describe each eigenspace.

(5)

- (a) What are the eigenvalues of the 2×2 zero matrix (the matrix all of whose entries are 0)? Describe each eigenspace.
- (b) Now let $n > 2$ be a positive integer. What are the eigenvalues of the $n \times n$ zero matrix? Describe each eigenspace.

(6) Label each of the following statements as True or False. Provide justification for your response.

- (a) **True/False** If $A\mathbf{v} = \lambda\mathbf{v}$, then λ is an eigenvalue of A with eigenvector \mathbf{v} .
- (b) **True/False** The scalar λ is an eigenvalue of a square matrix A if and only if the equation $(A - \lambda I_n)\mathbf{x} = \mathbf{0}$ has a nontrivial solution.
- (c) **True/False** If λ is an eigenvalue of a matrix A , then there is only one nonzero vector \mathbf{v} with $A\mathbf{v} = \lambda\mathbf{v}$.
- (d) **True/False** The eigenspace of an eigenvalue of an $n \times n$ matrix A is the same as $\text{Nul}(A - \lambda I_n)$.
- (e) **True/False** If \mathbf{v}_1 and \mathbf{v}_2 are eigenvectors of a matrix A corresponding to the same eigenvalue λ , then $\mathbf{v}_1 + \mathbf{v}_2$ is also an eigenvector of A .
- (f) **True/False** If \mathbf{v}_1 and \mathbf{v}_2 are eigenvectors of a matrix A , then $\mathbf{v}_1 + \mathbf{v}_2$ is also an eigenvector of A .
- (g) **True/False** If \mathbf{v} is an eigenvector of an invertible matrix A , then \mathbf{v} is also an eigenvector of A^{-1} .

Project: Modeling Population Migration

As introduced earlier, data from the Michigan Department of Technology, Management, and Budget shows that from 2011 to 2012, approximately 0.05% of the U.S. population outside of Michigan moved to the state of Michigan, while approximately 2% of Michigan's population moved out of Michigan. We are interested in determining the long-term distribution of population in Michigan.



Let $\mathbf{x}_n = \begin{bmatrix} m_n \\ u_n \end{bmatrix}$ be the 2×1 vector where m_n is the population of Michigan and u_n is the U.S. population outside of Michigan in year n . Assume that we start our analysis at generation 0 and $\mathbf{x}_0 = \begin{bmatrix} m_0 \\ u_0 \end{bmatrix}$.

Project Activity 14.1.

- (a) Explain how the data above shows that

$$m_1 = 0.98m_0 + 0.0005u_0$$

$$u_1 = 0.02m_0 + 0.9995u_0$$

- (b) Identify the matrix A such that $\mathbf{x}_1 = A\mathbf{x}_0$.

Once we have the equation $\mathbf{x}_1 = A\mathbf{x}_0$, we can extend it to subsequent years:

$$\mathbf{x}_2 = A\mathbf{x}_1, \quad \mathbf{x}_3 = A\mathbf{x}_2, \quad \dots, \quad \mathbf{x}_{n+1} = A\mathbf{x}_n$$

for each $n \geq 0$.

This example illustrates the general nature of what is called a *Markov process* (see Definition 9.4). Recall that the matrix A that provides the link from one generation to the next is called the transition matrix.

In situations like these, we are interested in determining if there is a steady-state vector, that is a vector that satisfies

$$\mathbf{x} = A\mathbf{x}. \tag{14.10}$$

Such a vector would show us the long-term population of Michigan provided the population dynamics do not change.

Project Activity 14.2.

- (a) Explain why a steady-state solution to (14.10) is an eigenvector of A . What is the corresponding eigenvalue?
- (b) Consider again the transition matrix A from Project Activity 14.1. Recall that the solutions to equation (14.10) are all the vectors in $\text{Nul}(A - I_2)$. In other words, the eigenvectors of A for this eigenvalue are the nonzero vectors in $\text{Nul}(A - I_2)$. Find a basis for the eigenspace of A corresponding to this eigenvalue. Use whatever technology is appropriate.
- (c) Once we know a basis for the eigenspace of the transition matrix A , we can use it to estimate the steady-state population of Michigan (assuming the stated migration trends are valid long-term). According to the US Census Bureau², the resident US population on December 1, 2019 was 330,073,471. Assuming no population growth in the U.S., what would the long-term population of Michigan be? How realistic do you think this is?

²<https://www.census.gov/data/tables/time-series/demo/popest/2010s-national-total.html>



Section 15

Bases and Dimension

Focus Questions

By the end of this section, you should be able to give precise and thorough answers to the questions listed below. You may want to keep these questions in mind to focus your thoughts as you complete the section.

- What is the dimension of a subspace of \mathbb{R}^n ? What property of bases makes the dimension a well-defined number?
- If W is a subspace of \mathbb{R}^n with dimension k , what must be true about any linearly independent subset S of W that contains exactly k vectors?
- If W is a subspace of \mathbb{R}^n with dimension k , what must be true about any subset S of W that contains exactly k vectors and spans W ?
- What is the rank of a matrix?
- What does the Rank-Nullity Theorem say?

Application: Lattice Based Cryptography

When you use your credit card, you expect that the information that is transmitted is protected so others can't use your card. Similarly, when you create a password for your computer or other devices, you do so with the intention that it will be difficult for others to decipher.

Cryptology is the study of methods to maintain secure communication in the presence of other parties (cryptography), along with the study of breaking codes (cryptanalysis). In essence, cryptology is the art of keeping and breaking secrets. The creation of secure codes (cryptography) can provide confidentiality (ensure that information is available only to the intended recipients), data integrity (prevent data from being altered between the sender and recipient), and authentication (making sure that the information is from the correct source).

Modern cryptology uses mathematical theory that can be implemented with computer hardware and algorithms. The security of public key systems is largely based on mathematical problems that

are very difficult to solve. For example, the security of the RSA system relies on the fact that it is computationally difficult to find prime factors of very large numbers, and elliptic curve cryptography relies on the difficulty of the discrete logarithm problem for elliptic curves. However, the continual increase in the power of computers threatens the security of these systems, and so cryptographic systems have to keep adapting to the newest technology. For example, Shor's Algorithm (which could run on a quantum computer) can solve the public key cryptographic systems that rely on the integer factorization problem or the discrete logarithm problem. So if a working quantum computer was ever developed, it would threaten the existing cryptographic systems. Lattice-based cryptography is a potential source of systems that may be secure even in such an environment. The security of these systems is dependent on the fact that the average case of the difficulty of certain problems in lattice theory is higher than the worst case problems that underpin current cryptosystems. As we will see later in this section, lattices are built on bases for subspace of \mathbb{R}^n .

Introduction

A basis provides a system in which we can uniquely represent every vector in the space we are considering. More specifically, every vector in the space can be expressed as a linear combination of the vectors in the basis in a unique way. In order to be able to cover every point in the space, the basis has to span the space. In order to be able to provide a unique coordinate for each point, there should not be any extra vectors in the basis, which is achieved by linear independence of the vectors. For practical reasons, a basis simplifies many problems because we only need to solve the problem for each of the basis vectors. Solutions of the other cases usually follow because every vector in the space can be expressed as a unique linear combination of the basis vectors.

Recall that a basis for a subspace W of \mathbb{R}^n is a set of vectors which are linearly independent and which span W .

Preview Activity 15.1.

- (1) For each of the following sets of vectors, determine whether the vectors form a basis of \mathbb{R}^3 . Use any appropriate technology for your computations.

$$(a) \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \right\}$$

$$(b) \left\{ \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 2 \\ 3 \\ 1 \end{bmatrix} \right\}$$

$$(c) \left\{ \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 3 \\ 3 \end{bmatrix}, \begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix} \right\}$$

$$(d) \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \right\}$$

- (2) In problem (1) we should have noticed that a space can have more than one basis, but that any two bases contain the same number of elements. This is a critically important idea that we



investigate in more detail in this problem in one specific case. Assume that W is a subspace of \mathbb{R}^n that has a basis $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2\}$ with two basis vectors. We want to see if any other basis for W can have a different number of elements. Let us now consider a set $U = \{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ of three vectors in W . Our goal is to determine if U can be a basis for W . Since \mathcal{B} is a basis for W , any vector in W can be written as a linear combination of the vectors in \mathcal{B} . So we can write

$$\mathbf{u}_1 = a_{11}\mathbf{v}_1 + a_{21}\mathbf{v}_2 \quad (15.1)$$

$$\mathbf{u}_2 = a_{12}\mathbf{v}_1 + a_{22}\mathbf{v}_2 \quad (15.2)$$

$$\mathbf{u}_3 = a_{13}\mathbf{v}_1 + a_{23}\mathbf{v}_2 \quad (15.3)$$

for some scalars a_{ij} . If U were to be a basis for W , then U would have to be a linearly independent set. To determine the independence or dependence of U we consider the vector equation

$$x_1\mathbf{u}_1 + x_2\mathbf{u}_2 + x_3\mathbf{u}_3 = \mathbf{0} \quad (15.4)$$

for scalars x_1, x_2 , and x_3 .

- (a) Substitute for \mathbf{u}_1 , \mathbf{u}_2 , and \mathbf{u}_3 from (15.1), (15.2), and (15.3) into (15.4) and perform some vector algebra to show that

$$\mathbf{0} = (a_{11}x_1 + a_{12}x_2 + a_{13}x_3)\mathbf{v}_1 + (a_{21}x_1 + a_{22}x_2 + a_{23}x_3)\mathbf{v}_2. \quad (15.5)$$

- (b) Recall that $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2\}$ is a basis. What does that tell us about the weights in the linear combination (15.5)? Explain why $A\mathbf{x} = \mathbf{0}$, where $A = [a_{ij}]$ and $\mathbf{x} = [x_1 \ x_2 \ x_3]^T$.
- (c) With A as in part (b), how many solutions does the system $A\mathbf{x} = \mathbf{0}$ have? Explain. (Hint: Consider the number of rows and columns of A .) What does this tell us about the independence or dependence of the set U ? Why?
- (d) Can U be a basis for W ? Explain.

The Dimension of a Subspace of \mathbb{R}^n

In Preview Activity 15.1 we saw that a subspace of \mathbb{R}^n can have more than one basis. This prompts the question of how, if at all, are any two bases for a given space related. More specifically, is it possible to have two bases for a given subspace of \mathbb{R}^n that contain different numbers of vectors? As we will see the answer is no, which will lead us to the concept of *dimension*.

Let W be a subspace of \mathbb{R}^n that has a basis $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k\}$ of k vectors. Since we have been calling bases minimal spanning sets, we should expect that any two bases for the same subspace have the same number of elements (otherwise one of the two bases would not be minimal). Our goal in this section is to prove that result – that any other basis of W contains exactly k vectors. The approach will be the same as was used in Preview Activity 15.1. We will let $U = \{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_m\}$ be a set of vectors in W with $m > k$ and demonstrate that U is a linearly dependent set. To argue linear dependence, let x_1, x_2, \dots, x_m be scalars so that

$$x_1\mathbf{u}_1 + x_2\mathbf{u}_2 + \dots + x_m\mathbf{u}_m = \mathbf{0}. \quad (15.6)$$



For each i there exist scalars a_{ij} so that

$$\mathbf{u}_i = a_{1i}\mathbf{v}_1 + a_{2i}\mathbf{v}_2 + \cdots + a_{ki}\mathbf{v}_k.$$

Substituting into (15.6) yields

$$\begin{aligned} \mathbf{0} &= x_1\mathbf{u}_1 + x_2\mathbf{u}_2 + \cdots + x_m\mathbf{u}_m \\ &= x_1(a_{11}\mathbf{v}_1 + a_{21}\mathbf{v}_2 + \cdots + a_{k1}\mathbf{v}_k) + x_2(a_{12}\mathbf{v}_1 + a_{22}\mathbf{v}_2 \\ &\quad + \cdots + a_{k2}\mathbf{v}_k) + \cdots + x_m(a_{1m}\mathbf{v}_1 + a_{2m}\mathbf{v}_2 + \cdots + a_{km}\mathbf{v}_k) \\ &= (x_1a_{11} + x_2a_{12} + x_3a_{13} + \cdots + x_ma_{1m})\mathbf{v}_1 \\ &\quad + (x_1a_{21} + x_2a_{22} + x_3a_{23} + \cdots + x_ma_{2m})\mathbf{v}_2 \\ &\quad + \cdots + (x_1a_{k1} + x_2a_{k2} + x_3a_{k3} + \cdots + x_ma_{km})\mathbf{v}_k. \end{aligned} \tag{15.7}$$

Since \mathcal{B} is a basis, the vectors $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k$ are linearly independent. So each coefficient in (15.7) is 0 and $\mathbf{x} = [x_1 \ x_2 \ \cdots \ x_m]^\top$ is a solution to the homogeneous system $A\mathbf{x} = \mathbf{0}$, where $A = [a_{ij}]$. Now A is a $k \times m$ matrix with $m > k$, so not every column of A is a pivot column. This means that $A\mathbf{x} = \mathbf{0}$ has a nontrivial solution. It follows that the vector equation (15.6) has a nontrivial solution and so the m vectors $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_m$ are linearly dependent. We summarize this in the following theorem.

Theorem 15.1. *Let W be a subspace of \mathbb{R}^n containing a basis with k vectors. If $m > k$, then any set of m vectors in W is linearly dependent.*

One consequence of Theorem 15.1 is that, in addition to being a minimal spanning set, a basis is also a maximal linearly independent set.

Activity 15.1. Now let's return to the question of the number of elements in a basis for a subspace of \mathbb{R}^n . Recall that we are assuming that W has a basis $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k\}$ of k vectors in \mathbb{R}^n . Suppose that \mathcal{B}' is another basis for W containing m vectors.

- (a) Given the fact that \mathcal{B} is a basis for W , what does Theorem 15.1 tell us about the relationship between m and k ?
- (b) Given the fact that \mathcal{B}' is a basis for W , what does Theorem 15.1 tell us about the relationship between m and k ?
- (c) What do the results of (a) and (b) tell us about the relationship between m and k ? What can we conclude about any basis for W ?

The result of Activity 15.1 is summarized in the following theorem. Recall that the trivial space is the single element set $\{\mathbf{0}\}$.

Theorem 15.2. *If a nontrivial subspace W of \mathbb{R}^n has a basis of k vectors, then every basis of W contains exactly k vectors.*

This last theorem states that the number of vectors in a basis for a subspace space is a well-defined number. In other words, the number of vectors in a basis is an *invariant* of the subspace. This important number is given a name.



Definition 15.3. The **dimension** of a subspace W of \mathbb{R}^n is the number of vectors in a basis for W . The dimension of the trivial subspace $\{\mathbf{0}\}$ of \mathbb{R}^n is defined to be 0.

We denote the dimension of a subspace W of \mathbb{R}^n by $\dim(W)$. As we will see later, any two vector spaces of the same dimension are basically the same vector space. So the dimension of a vector space is an important number that essentially tells us all we need to know about the structure of the space.

Activity 15.2. Find the dimensions of each of the indicated subspaces of \mathbb{R}^n for the appropriate n . Explain your method.

(a) $\text{Span} \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 2 \\ 3 \\ 0 \end{bmatrix} \right\}$

(b) xy -plane in \mathbb{R}^3

(c) \mathbb{R}^3

(d) \mathbb{R}^n

Conditions for a Basis of a Subspace of \mathbb{R}^n

There are two items we need to confirm before we can state that a subset \mathcal{B} of a subspace W of \mathbb{R}^n is a basis for W : the set \mathcal{B} must be linearly independent and span W . However, if we have the right number (namely, the dimension) of vectors in our set \mathcal{B} , then either one of these conditions will imply the other.

Activity 15.3. Let W be a subspace of \mathbb{R}^n with $\dim(W) = k$. We know that every basis of W contains exactly k vectors.

- (a) Suppose that S is a subset of W that contains k vectors and is linearly independent. In this part of the activity we will show that S must span W .
 - i. Suppose that S does not span W . Explain why this implies that W contains a set of $k + 1$ linearly independent vectors.
 - ii. Explain why the result of part i. tells us that S is a basis for W .
- (b) Now suppose that S is a subset of W with k vectors that spans W . In this part of the activity we will show that S must be linearly independent.
 - i. Suppose that S is not linearly independent. Explain why we can then find a proper subset of S that is linearly independent but has the same span as S .
 - ii. Explain why the result of part i. tells us that S is a basis for W .

The result of Activity 15.3 is the following important theorem.

Theorem 15.4. Let W be a subspace of \mathbb{R}^n of dimension k and let S be a subset of W containing exactly k vectors.



- (1) If S is linearly independent, then S is a basis for W .
- (2) If S spans W , then S is a basis for W .

Finding a Basis for a Subspace

Since every vector in a subspace of \mathbb{R}^n can be written uniquely as a linear combination of vectors in a basis for the subspace, a basis provides us with the most efficient and convenient way to represent vectors in the subspace. Until now we have been given a set of vectors and have been asked to find a basis from that set, so an important question to address is how we can find a basis for a subspace W of \mathbb{R}^n starting from scratch. Here is one way. If $W = \{\mathbf{0}\}$, then the dimension of W is 0 and W has no basis. So suppose $\dim(W) > 0$. Start by choosing any nonzero vector \mathbf{w}_1 in W . Let $\mathcal{B}_1 = \{\mathbf{w}_1\}$. If \mathcal{B}_1 spans W , then \mathcal{B}_1 is a basis for W . If not, there is a vector \mathbf{w}_2 in W that is not in $\text{Span}(\mathcal{B}_1)$. Then $\mathcal{B}_2 = \{\mathbf{w}_1, \mathbf{w}_2\}$ is a linearly independent set. If $\text{Span}(\mathcal{B}_2) = W$, then \mathcal{B}_2 is a basis for W and we are done. If not, repeat the process. Since any basis for W can contain at most $n = \dim(\mathbb{R}^n)$ vectors, we know the process must stop at some point. This process also allows us to construct a basis for a vector space that contains a given nonzero vector.

Activity 15.4. Find a basis for \mathbb{R}^3 that contains the vector $\begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}$. When constructing your basis, how do you know when to stop?

Rank of a Matrix

In this section, we define the rank of a matrix and review conditions to add to our Invertible Matrix Theorem.

Activity 15.5. Let $A = \begin{bmatrix} 1 & 2 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & 1 & 1 \end{bmatrix}$.

- (a) Without performing any calculations, find $\dim(\text{Nul } A)$. Explain.
- (b) Without performing any calculations, find $\dim(\text{Col } A)$. Explain.
- (c) There is a connection between $\dim(\text{Nul } A)$, $\dim(\text{Col } A)$ and the size of A . Find this connection and explain it.

As Activity 15.5 illustrates, the number of vectors in a basis for $\text{Nul } A$ is the number of non-pivot columns in A and the number of vectors in a basis for $\text{Col } A$ is the number of pivot columns of A . We define the *rank* of a matrix A (denoted $\text{rank}(A)$) to be the dimension of $\text{Col } A$ and the *nullity* of A to be dimension of $\text{Nul } A$. The dimension of the null space of A is also called the *nullity* of A (denoted $\text{nullity}(A)$). Using this terminology we have the Rank-Nullity Theorem.

Theorem 15.5 (The Rank-Nullity Theorem). *Let A be an $m \times n$ matrix. Then*

$$\text{rank}(A) + \text{nullity}(A) = n.$$



There is also a row space of a matrix A , which we define to be the span of the rows of A . We can find the row space of A by finding the column space of A^T , so the row space is really nothing new. As it turns out, the dimension of the row space of A is always equal to the dimension of the column space of A , and justification for this statement is in the exercises.

The Rank-Nullity Theorem allows us to add extra conditions to the Invertible Matrix Theorem.

Theorem 15.6 (The Invertible Matrix Theorem). *Let A be an $n \times n$ matrix. The following statements are equivalent.*

- (a) *The matrix A is an invertible matrix.*
- (b) *The matrix equation $A\mathbf{x} = \mathbf{0}$ has only the trivial solution.*
- (c) *The matrix A has n pivot columns.*
- (d) *Every row of A contains a pivot.*
- (e) *The columns of A span \mathbb{R}^n .*
- (f) *The matrix A is row equivalent to the identity matrix I_n .*
- (g) *The columns of A are linearly independent.*
- (h) *The columns of A form a basis for \mathbb{R}^n .*
- (i) *The matrix transformation T from \mathbb{R}^n to \mathbb{R}^n defined by $T(\mathbf{x}) = A\mathbf{x}$ is one-to-one.*
- (j) *The matrix equation $A\mathbf{x} = \mathbf{b}$ has exactly one solution for each vector \mathbf{b} in \mathbb{R}^n .*
- (k) *The matrix transformation T from \mathbb{R}^n to \mathbb{R}^n defined by $T(\mathbf{x}) = A\mathbf{x}$ is onto.*
- (l) *There is an $n \times n$ matrix C so that $AC = I_n$.*
- (m) *There is an $n \times n$ matrix D so that $DA = I_n$.*
- (n) *The scalar 0 is not an eigenvalue of A .*
- (o) *The matrix A^T is invertible.*
- (p) *$\text{Nul } A = \{\mathbf{0}\}$.*
- (q) *$\text{Col } A = \mathbb{R}^n$.*
- (r) *$\dim(\text{Col } A) = n$*
- (s) *$\dim(\text{Nul } A) = 0$*
- (t) *$\text{rank}(A) = n$*

Examples

What follows are worked examples that use the concepts from this section.



Example 15.7. Let $W = \left\{ \begin{bmatrix} r+s+u \\ r+3s+2t-u \\ -s-t+u \\ s+t-u \end{bmatrix} : r, s, t, u \in \mathbb{R} \right\}$.

- (a) Explain why W is a subspace of \mathbb{R}^4 .
- (b) Find a basis for W and determine the dimension of W .

Example Solution.

- (a) We can write any vector in W in the form

$$\begin{bmatrix} r+s+u \\ r+3s+2t-u \\ -s-t+u \\ s+t-u \end{bmatrix} = r \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix} + s \begin{bmatrix} 1 \\ 3 \\ -1 \\ 1 \end{bmatrix} + t \begin{bmatrix} 0 \\ 2 \\ -1 \\ 1 \end{bmatrix} + u \begin{bmatrix} 1 \\ -1 \\ 1 \\ -1 \end{bmatrix},$$

so

$$W = \text{Span} \left\{ \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 3 \\ -1 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 2 \\ -1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \\ 1 \\ -1 \end{bmatrix} \right\}.$$

As a span of a set of vectors in \mathbb{R}^4 , W is a subspace of \mathbb{R}^4 .

(b) Let $A = \begin{bmatrix} 1 & 1 & 0 & 1 \\ 1 & 3 & 2 & -1 \\ 0 & -1 & -1 & 1 \\ 0 & 1 & 1 & -1 \end{bmatrix}$. To find a basis for W , we note that the reduced row echelon form of A is $\begin{bmatrix} 1 & 0 & -1 & 2 \\ 0 & 1 & 1 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$. Since the pivot columns of A form a basis for $\text{Col } A = W$, we conclude that

$$\left\{ \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 3 \\ -1 \\ 1 \end{bmatrix} \right\}$$

is a basis for W . Therefore, $\dim(W) = 2$.

Example 15.8. Find a basis and the dimension of the solution set to the system

$$\begin{aligned} r + s - t + 2u &= 0 \\ 3r - s + 2t - u &= 0 \\ r - 3s + 4t - 5u &= 0 \\ 5r - 3s + 5t - 4u &= 0. \end{aligned}$$



Example Solution.

The coefficient matrix of this system is

$$A = \begin{bmatrix} 1 & 1 & -1 & 2 \\ 3 & -1 & 2 & -1 \\ 1 & -3 & 4 & -5 \\ 5 & -3 & 5 & -4 \end{bmatrix},$$

and the solution set to the system is $\text{Nul } A$. To find a basis for $\text{Nul } A$ we row reduce A to

$$\begin{bmatrix} 1 & 0 & \frac{1}{4} & \frac{1}{4} \\ 0 & 1 & -\frac{5}{4} & \frac{7}{4} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

The general solution to the system has the form

$$\begin{bmatrix} r \\ s \\ t \\ u \end{bmatrix} = \begin{bmatrix} -\frac{1}{4}t - \frac{1}{4}u \\ \frac{5}{4}t - \frac{7}{4}u \\ t \\ u \end{bmatrix} = t \begin{bmatrix} -\frac{1}{4} \\ \frac{5}{4} \\ 1 \\ 0 \end{bmatrix} + u \begin{bmatrix} -\frac{1}{4} \\ -\frac{7}{4} \\ 0 \\ 1 \end{bmatrix},$$

so $\left\{ \begin{bmatrix} -\frac{1}{4} \\ \frac{5}{4} \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -\frac{1}{4} \\ -\frac{7}{4} \\ 0 \\ 1 \end{bmatrix} \right\}$ is a basis for $\text{Nul } A$ and $\dim(\text{Nul } A) = 2$.

Summary

The key idea in this section is the *dimension* of a vector space.

- Any two bases for a vector space *must* contain the same number of vectors. Therefore, we can define the *dimension* of a vector space W to be the number of vectors in any basis for W .
- If W is a subspace of \mathbb{R}^n with dimension k and S is any linearly independent subset of W with k vectors, then S is a basis for W .
- If W is a subspace of \mathbb{R}^n with dimension k and S is any subset of W with k vectors that spans W , then S is a basis for W .
- The rank of a matrix is the dimension of its column space.
- The Rank-Nullity Theorem states that if A is an $m \times n$ matrix, then $\text{rank}(A) + \text{nullity}(A) = n$.

Exercises

(1) Let $A = \begin{bmatrix} 1 & 3 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 2 & 6 & 0 & 0 & 1 \\ 1 & 3 & 2 & 4 & 1 \\ 3 & 9 & 1 & 2 & -1 \\ 3 & 9 & 3 & 6 & 1 \end{bmatrix}$.

(a) Find a basis for $\text{Col } A$. What is the dimension of $\text{Col } A$? What, then, is the dimension of $\text{Nul } A$?

(b) Find a basis for $\text{Nul } A$ and verify the dimension you found in part (a).

(2) Let $A = \begin{bmatrix} 2 & -1 & 1 \\ 1 & 0 & 1 \\ 1 & -1 & 2 \end{bmatrix}$. The eigenvalues of A are 1 and 2. Find the dimension of each eigenspace of A .

(3) Let $A = \begin{bmatrix} 1 & 2 & -1 & -1 \\ -2 & -4 & 2 & 2 \\ 1 & 2 & -1 & -1 \end{bmatrix}$.

(a) Find a basis for $\text{Col } A$. What is the rank of A ?

(b) Find a basis for $\text{Nul } A$. What is the nullity of A ?

(c) Verify the Rank-Nullity Theorem for A .

(d) The row space of A is the span of the rows of A . Find a basis for the row space of A and the dimension of the row space of A .

(4) Let A be an $m \times n$ matrix with r pivots, where r is less than or equal to both m, n . Fill in the blanks.

(a) The null space of A is a subspace of _____.

(b) The column space of A is a subspace of _____.

(c) Suppose $r = m$. Then there is a pivot in every _____ and $\text{Col } A = \text{_____}$.

(d) Suppose $r = n$. Then there is a pivot in every _____ and $\text{Nul } A = \text{_____}$.

(e) If A has 3 pivots, then the rank of A is _____.

(f) If A has 3 pivots, then the number of free variables in the system $Ax = \mathbf{0}$ is _____.

(g) The dimension of $\text{Col } A$ is equal to the number of _____, i.e. $\dim \text{Col } A = \text{_____}$.

(h) The dimension of $\text{Nul } A$ is equal to the number of _____, i.e. $\dim \text{Nul } A = \text{_____}$.

(i) $\dim(\text{Nul } A) + \dim(\text{Col } A) = \text{_____}$.



- (j) Suppose the columns of A span \mathbb{R}^m . Then rank A is _____.
- (k) Suppose the columns of A are linearly independent. Then $r =$ _____ and the dimension of $\text{Nul } A$ is _____.
- (5) Prove the remaining parts of the Invertible Matrix Theorem (Theorem 15.6). Let A be an $n \times n$ matrix.
- Prove that A is invertible if and only if $\dim(\text{Nul } A) = 0$.
 - Prove that A is invertible if and only if $\dim(\text{Col } A) = n$.
- (6) We can convert the language of the Rank-Nullity Theorem to matrix transformation language, as we show in this exercise. Let T be the matrix transformation defined by the matrix A .
- How is the kernel of T related to A ?
 - How is the range of T related to A ?
 - How is the domain of T related to A ?
 - Explain why the Rank-Nullity Theorem says that $\dim(\text{Ker}(T)) + \dim(\text{Range}(T)) = \dim(\text{Domain}(T))$.
- (7) Let W be a subspace of \mathbb{R}^4 . What are possible values for the dimension of W ? Explain. What are the geometric descriptions of W in each case?
- (8) Is it possible to find two subspaces W_1 and W_2 in \mathbb{R}^3 such that $W_1 \cap W_2 = \{\mathbf{0}\}$ and $\dim W_1 = \dim W_2 = 2$? If possible, give an example and justify that they satisfy the conditions. If not possible, explain why not. (Hint: Dimension two leads to two linearly independent vectors in each of W_i .)
- (9) Determine the dimensions of the column space and null space of
- $$\begin{bmatrix} 1 & 2 & 4 & 3 & 2 \\ 1 & 0 & 2 & 1 & 4 \\ 1 & 1 & 3 & 1 & 2 \\ 1 & 0 & 2 & 2 & 5 \end{bmatrix}.$$
- (10) If possible, find a 3×4 matrix whose column space has dimension 3 and null space has dimension 1. Explain how you found the matrix in addition to explaining why your answer works. If not possible, explain why it is not possible to find such a matrix.
- (11)
- If possible, find a 5×5 matrix whose column space has the same dimension as its null space. Explain how you found the matrix in addition to explaining why your answer works. If not possible, explain why it is not possible to find such a matrix.
 - If possible, find a matrix A so that $\text{Col } A = \text{Nul } A$. Explain how you found the matrix in addition to explaining why your answer works. If not possible, explain why it is not possible to find such a matrix.
- (12) In this exercise we examine why the dimension of a row space of a matrix is the same as the dimension of the column space of the matrix. Let A be an $m \times n$ matrix.

- (a) Explain why row operations do not change the row space of a matrix. Then explain why if R is the reduced row echelon form of A , then $\text{Row } R = \text{Row } A$, where $\text{Row } M$ is the row space of the matrix M .
- (b) Explain why the rows of R that contain pivots form a basis for $\text{Row } R$, and also of $\text{Row } A$.
- (c) Explain why $\text{rank}(A)$ is the number of pivots in the matrix A . Then explain why $\dim(\text{Row } A) = \dim(\text{Col } A)$.
- (13) Label each of the following statements as True or False. Provide justification for your response.
- True/False** The dimension of the column space of a 3×2 matrix can be three.
 - True/False** There exists a 3×3 matrix whose column space has equal dimension as the null space.
 - True/False** If a set of vectors spans a subspace, then that set is a basis of this subspace.
 - True/False** If a linearly independent set of vectors spans a subspace, then that set is a basis of this subspace.
 - True/False** The dimension of a space is the minimum number of vectors needed to span that space.
 - True/False** The dimension of the null space of a 3×2 matrix can at most be 2.
 - True/False** Any basis of \mathbb{R}^4 contains 4 vectors.
 - True/False** If n vectors span \mathbb{R}^n , then these vectors form a basis of \mathbb{R}^n .
 - True/False** Every line in \mathbb{R}^n is a one-dimensional subspace of \mathbb{R}^n .
 - True/False** Every plane through origin in \mathbb{R}^n is a two-dimensional subspace of \mathbb{R}^n .
 - True/False** In \mathbb{R}^n any n linearly independent vectors form a basis.

Project: The GGH Cryptosystem

A cryptographic system (or cryptosystem) allows for secure communication between two or more parties. These systems take messages (called *plaintext*) and encrypt them in some way to produce what is called *ciphertext*. This is the scrambled information that is transmitted to the receiver, from which it should not be possible for someone who does not have the proper key to recover the original message. When the message is received by the intended recipient, it must be unscrambled or *decrypted*. Decryption is the process of converting ciphertext back to plaintext.

The Goldreich-Goldwasser-Halevi (GGH) public key cryptosystem¹ uses lattices to encrypt plaintext. The security of the system depends on the fact that the Closest Vector Problem (CVP) is, in general, a very hard problem. To begin to understand these cryptosystems, we begin with lattices.

¹Published in 1997 by Oded Goldreich, Shafi Goldwasser, and Shai Halevi.



Lattices are closely related to spans of sets of vectors in \mathbb{R}^n . If we start with a linearly independent set $S = \{\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_m\}$ in \mathbb{R}^n , we can create the span of S – the set of all linear combinations

$$c_1\mathbf{b}_1 + c_2\mathbf{b}_2 + \cdots + c_m\mathbf{b}_m,$$

where c_1, c_2, \dots, c_m are real numbers. This span creates a subspace of \mathbb{R}^n . If we restrict the set from which we choose the coefficients, we can create different types of structures. An important one is a lattice. The *lattice* $\mathcal{L}(S)$ defined by the linearly independent set $S = \{\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_m\}$ is the set of linear combinations

$$c_1\mathbf{b}_1 + c_2\mathbf{b}_2 + \cdots + c_m\mathbf{b}_m,$$

where c_1, c_2, \dots, c_m are integers. If the vectors in S have integer components, then every point in $\mathcal{L}(S)$ will have integer entries. In these cases, $\mathcal{L}(S)$ is a subset of \mathbb{Z}^n , as illustrated in Figure 15.1. Also, if $m = n$ we say that the lattice $\mathcal{L}(S)$ is *full-rank*. We will restrict ourselves to full-rank lattices in this project. A *basis* for a lattice is any set of linearly independent vectors that generates the lattice. There is a little special notation that is often used with lattices. If $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_n\}$ is a basis for \mathbb{R}^n , we associate to \mathcal{B} the matrix $B = [\mathbf{b}_1 \ \mathbf{b}_2 \ \mathbf{b}_3 \ \cdots \ \mathbf{b}_n]$. We then use the notation $\mathcal{L}(B)$ to also refer to the lattice defined by \mathcal{B} .

Project Activity 15.1. We explore lattices in more detail in this activity.

- (a) Let $S_1 = \{[1 1]^\top, [-1 1]^\top\}$.

- i. Find five distinct vectors in $\mathcal{L}(S_1)$.
- ii. Is the vector $[1 0]^\top$ in $\mathcal{L}(S_1)$? Justify your answer.
- iii. We can draw pictures of lattices by plotting the terminal points of the lattice vectors.
Draw all of the lattice points in $\mathcal{L}(S_1)$ on the square with vertices $(-4, -4)$, $(4, -4)$, $(4, 4)$, and $(-4, 4)$.

- (b) Now let $S_2 = \{[3 5]^\top, [1 2]^\top\}$. A picture of $\mathcal{L}(S_2)$ is shown in Figure 15.1 with the basis vectors highlighted. As we have seen, $\mathcal{L}(S_1)$ is not the entire space \mathbb{Z}^2 . Is $\mathcal{L}(S_2) = \mathbb{Z}^2$? Justify your answer.

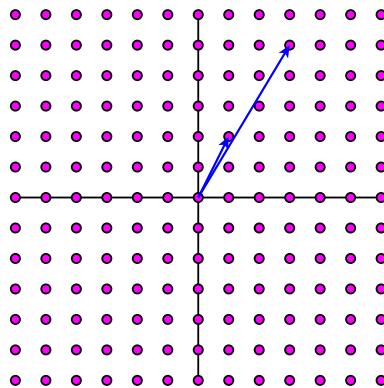


Figure 15.1: The lattice $\mathcal{L}(S_2)$.

Project Activity 15.1 shows that even if \mathcal{B} is a basis for \mathbb{R}^n , it does not follow that $\mathcal{L}(\mathcal{B})$ is all of \mathbb{Z}^n . So lattices can be complicated, and problems in lattice theory can be very difficult.

The GGH cryptosystem relies on the fact that we can convert “good” bases for lattices into “bad” bases. We will not delve into the details of what separates a “good” basis from a “bad” one, but suffice it to say that a good basis is one in which the basis vectors are close to being perpendicular² and are all short (that is, they have small norms), while any other basis is a bad basis. An example of a good basis is the basis S_1 for \mathbb{R}^2 in Project Activity 15.1, and we will see later that $\{[-2 \ 8]^\top, [-1 \ 3]^\top\}$ is a bad basis for the same lattice. You should draw a picture of the vectors $[-2 \ 8]^\top$ and $[-1 \ 3]^\top$ to convince yourself that this is a bad basis for its lattice.

The GGH cryptosystem works with two *keys* – a public key and a private key. The keys are based on lattices. The general process of the GGH cryptosystem is as follows. Begin with a good basis $\mathcal{B} = \{\mathbf{b}_1 \ \mathbf{b}_2 \ \dots \ \mathbf{b}_n\}$ of \mathbb{R}^n of vectors with integer components. Let $B = [\mathbf{b}_1 \ \mathbf{b}_2 \ \dots \ \mathbf{b}_n]$ be the matrix associated with \mathcal{B} . Let $\mathcal{B}' = \{\mathbf{b}'_1, \mathbf{b}'_2, \dots, \mathbf{b}'_n\}$ be a bad basis for which $\mathcal{L}(\mathcal{B}') = \mathcal{L}(\mathcal{B})$. Let $B' = [\mathbf{b}'_1 \ \mathbf{b}'_2 \ \dots \ \mathbf{b}'_n]$ be the matrix associated to the basis \mathcal{B}' . The bad basis can be shared with anyone (the public key), but the good basis is kept secret (the private key). Start with a message $\mathbf{m} = [m_1 \ m_2 \ \dots \ m_n]^\top$ with integer entries to send.

First we encrypt the message, which can be done by anyone who has the public key \mathcal{B}' .

- Create the message vector

$$\mathbf{m}' = m_1 \mathbf{b}'_1 + m_2 \mathbf{b}'_2 + \dots + m_n \mathbf{b}'_n = B' \mathbf{m}$$

that is in the lattice using the bad basis \mathcal{B}' .

- Choose a small error \mathbf{e} to add to \mathbf{m}' to move \mathbf{m}' off the lattice (small enough so that \mathbf{m}' does not pass by another lattice point). This is an important step that will make the message difficult to decrypt without the key. Let $\mathbf{c} = \mathbf{m}' + \mathbf{e} = B' \mathbf{m} + \mathbf{e}$. The vector \mathbf{c} is the ciphertext that is to be transmitted to the receiver.

Only someone who knows the basis \mathcal{B} can decode the ciphertext. This is done as follows.

- Find the vector $\mathbf{a} = a_1 \mathbf{b}_1 + a_2 \mathbf{b}_2 + \dots + a_n \mathbf{b}_n$ in the good basis \mathcal{B} that is closest to \mathbf{c} .
- We interpret the vector $[a_1 \ a_2 \ \dots \ a_n]^\top$ as being the encoded vector without the error. So to recreate the original message vector we need to undo the encrypting using the bad basis \mathcal{B}' . That is, we need to find the weights y_1, y_2, \dots, y_n such that

$$[a_1 \ a_2 \ \dots \ a_n]^\top = y_1 \mathbf{b}'_1 + y_2 \mathbf{b}'_2 + \dots + y_n \mathbf{b}'_n = B' [y_1 \ y_2 \ \dots \ y_n]^\top.$$

We can do this by as $[y_1 \ y_2 \ \dots \ y_n]^\top = B'^{-1} [a_1 \ a_2 \ \dots \ a_n]^\top$.

There are several items to address before we can implement this algorithm. One is how we create a bad basis \mathcal{B}' from \mathcal{B} that produces the same lattice. Another is how we find the vector in \mathcal{B} closest to a given vector. The latter problem is called the Closest Vector Problem (CVP) and is, in general, a very difficult problem. This is what makes lattice-based cryptosystems secure. We address the first of these items in the next activity, and the second a bit later.

²This is also a good property in vector spaces. We will see in a later section that perpendicular basis vectors make calculations in vector spaces relatively easy. A similar thing is true in lattices, where we are able to solve certain variants of closest vector problem very efficiently.



Project Activity 15.2. Consider again the basis $S_1 = \{[1 1]^\top, [-1 1]^\top\}$ from Project Activity 15.1, and let $B = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$ be the matrix whose columns are the vectors in S_1 .

- (a) Let T be the triangle with vertices $(0, 0)$, $(1, 1)$, and $(-1, 1)$. Show that this triangle is a right triangle, and conclude that the vectors $\mathbf{v}_1 = [1 1]^\top$, and $\mathbf{v}_2 = [-1 1]^\top$ are perpendicular.
- (b) Let $U = \begin{bmatrix} 3 & 1 \\ 5 & 2 \end{bmatrix}$. Let S_3 be the set whose vectors are the columns of the matrix $B_1 U$. Show that $\mathcal{L}(S_1) = \mathcal{L}(S_3)$.

The two bases $S_1 = \{[1 1]^\top, [-1 1]^\top\}$ and $S_3 = \{[-2 8]^\top, [-1 3]^\top\}$ from Project Activity 15.2 are shown in Figure 15.2. This figure illustrates how the matrix U transforms the basis S_1 , in which the vectors are perpendicular and short, to one in which the vectors are nearly parallel and significantly longer. So the matrix U converts the “good” basis S_1 into a “bad” basis S_3 , keeping the lattice intact. This is a key idea in the GGH cryptosystem. What makes this work is the fact that both U and U^{-1} have integer entries. The reason for this is that, for a 2×2 matrix $U = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$,

we know that $U^{-1} = \frac{1}{ad-bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$. If U has integer entries and $ad - bc = \pm 1$, then U^{-1} will also have integer entries. The number $ad - bc$ is called the *determinant* of U , and matrices with determinant of 1 or -1 are called *unimodular*. That what happened in Project Activity 15.2 happens in the general case is the focus of the next activity.

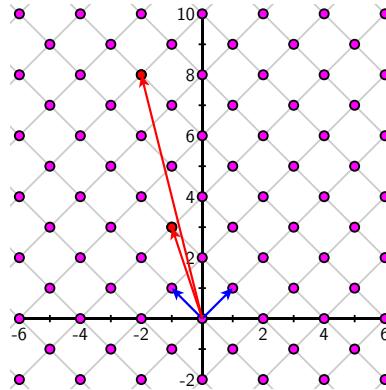


Figure 15.2: The lattices $\mathcal{L}(S_1)$ and $\mathcal{L}(S_3)$.

Project Activity 15.3. We will restrict ourselves to 2×2 matrices in this activity, but the results generalize to $n \times n$ matrices. Let $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2\}$ and $\mathcal{B}' = \{\mathbf{b}'_1, \mathbf{b}'_2\}$ be bases for \mathbb{R}^2 with integer entries, and let $B = [\mathbf{b}_1 \ \mathbf{b}_2]$ and $B' = [\mathbf{b}'_1 \ \mathbf{b}'_2]$ be the matrices associated to these bases. Show that if $B' = BU$ for some unimodular matrix U with integer entries, then $\mathcal{L}(B) = \mathcal{L}(B')$.

Project Activity 15.3 is the part we need for our lattice-based cryptosystem. Although we won’t show it here, the converse of the statement in Project Activity 15.3 is also true. That is, if \mathcal{B} and \mathcal{B}' generate the same lattice, then $B' = BU$ for some unimodular matrix U with integer entries.

There is one more item to address before we implement the GGH cryptosystem. That item is how to solve the Closest Vector Problem. There are some algorithms for approximating the closest vector in a basis. One is Babai's Closest Vector algorithm. This algorithm works in the following way. Consider a lattice with basis $\{\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_n\}$. To approximate the closest vector in the lattice to a vector \mathbf{w} , find the weights c_1, c_2, \dots, c_n in \mathbb{R} such that $\mathbf{w} = c_1\mathbf{b}_1 + c_2\mathbf{b}_2 + \dots + c_n\mathbf{b}_n$. Then round the coefficients to the nearest integer. This algorithm works well for a good basis, but is unlikely to return a lattice point that is close to \mathbf{w} if the basis is a bad one.

Now we put this all together to illustrate the GGH algorithm.

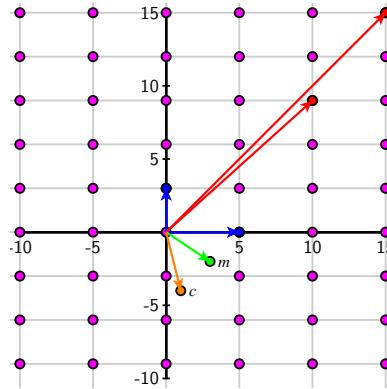


Figure 15.3: Decrypting an encrypted message.

Project Activity 15.4. Let $\mathcal{B} = \{[5 \ 0]^\top, [0 \ 3]^\top\}$ be the private key, and let $B = \begin{bmatrix} 5 & 0 \\ 0 & 3 \end{bmatrix}$ be the matrix whose columns are the vectors in \mathcal{B} . Let U be the unimodular matrix $U = \begin{bmatrix} 2 & 3 \\ 3 & 5 \end{bmatrix}$. Let $\mathbf{m} = [3 \ -2]^\top$ be our message and let $\mathbf{e} = [1 \ -1]^\top$ be our error vector.

- Use the unimodular matrix U to create the bad basis \mathcal{B}' .
- Determine the ciphertext message \mathbf{c} .
- A picture of the message vector \mathbf{m} and the ciphertext vector \mathbf{c} are shown in Figure 15.3. Although the closest vector in the lattice to \mathbf{c} can be determined by the figure, actual messages are constructed in high dimensional spaces where a visual approach is not practical. Use Babai's algorithm to find the vector in $\mathcal{L}(\mathcal{B})$ that is closest to \mathbf{c} and compare to Figure 15.3.
- The final step in the GGH scheme is to recover the original message. Complete the GGH algorithm to find this message.
- The GGH cryptosystem works because the CVP can be reasonably solved using a good basis. That is, Babai's algorithm works if our basis is a good basis. To illustrate that a bad basis will not allow us to reproduce the original message vector, show that Babai's algorithm does not return the closest vector to \mathbf{c} using the bad basis \mathcal{B}' .



Part IV

Eigenvalues and Eigenvectors

Section 16

The Determinant

Focus Questions

By the end of this section, you should be able to give precise and thorough answers to the questions listed below. You may want to keep these questions in mind to focus your thoughts as you complete the section.

- How do we calculate the determinant of an $n \times n$ matrix?
- What is one important fact the determinant tells us about a matrix?

Application: Area and Volume

Consider the problem of finding the area of a parallelogram determined by two vectors \mathbf{u} and \mathbf{v} , as illustrated at left in Figure 16.1. We could calculate this area, for example, by breaking up

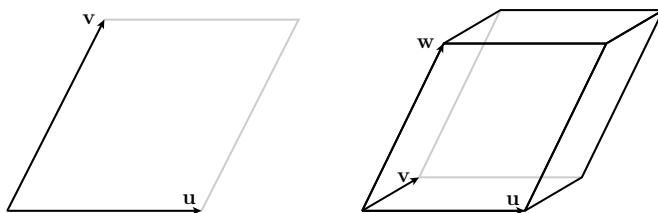


Figure 16.1: A parallelogram and a parallelepiped.

the parallelogram into two triangles and a rectangle and finding the area of each. Now consider the problem of calculating the volume of the three-dimensional analog (called a *parallelepiped*) determined by three vectors \mathbf{u} , \mathbf{v} , and \mathbf{w} as illustrated at right in Figure 16.1.

It is quite a bit more difficult to break this parallelepiped into subregions whose volumes are easy to compute. However, all of these computations can be made quickly by using determinants. The details are later in this section.

Introduction

We know that a non-zero vector \mathbf{x} is an eigenvector of an $n \times n$ matrix A if $A\mathbf{x} = \lambda\mathbf{x}$ for some scalar λ . Note that this equation can be written as $(A - \lambda I_n)\mathbf{x} = \mathbf{0}$. Until now, we were given eigenvalues of matrices and have used the eigenvalues to find the eigenvectors. In this section we will learn an algebraic technique to find the eigenvalues ourselves. We will also be able to justify why an $n \times n$ matrix has at most n eigenvalues.

A scalar λ is an eigenvalue of A if $(A - \lambda I_n)\mathbf{x} = \mathbf{0}$ has a non-trivial solution \mathbf{x} , which happens if and only if $A - \lambda I_n$ is not invertible. In this section we will find a scalar whose value will tell us when a matrix is invertible and when it is not, and use this scalar to find the eigenvalues of a matrix.

Preview Activity 16.1. In this activity, we will focus on 2×2 matrices. Let $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ be a 2×2 matrix. To see if A is invertible, we row reduce A by replacing row 2 with $a \cdot (\text{row 2}) - c \cdot (\text{row 1})$:

$$\begin{bmatrix} a & b \\ 0 & ad - bc \end{bmatrix}.$$

So the only way A can be reduced I_2 is if $ad - bc \neq 0$. We call this quantity $ad - bc$ the *determinant* of A , and denote the determinant of A as $\det(A)$ or $|A|$. When $\det(A) \neq 0$, we know that

$$A^{-1} = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}.$$

We now consider how we can use the determinant to find eigenvalues and other information about the invertibility of a matrix.

- (1) Let $A = \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}$. Find $\det(A)$ by hand. What does this mean about the matrix A ? Can you confirm this with other methods?
- (2) One of the eigenvalues of $A = \begin{bmatrix} 1 & 3 \\ 2 & 2 \end{bmatrix}$ is $\lambda = 4$. Recall that we can rewrite the matrix equation $A\mathbf{x} = 4\mathbf{x}$ in the form $(A - 4I_2)\mathbf{x} = \mathbf{0}$. What must be true about $A - 4I_2$ in order for 4 to be an eigenvalue of A ? How does this relate to $\det(A - 4I_2)$?
- (3) Another eigenvalue of $A = \begin{bmatrix} 1 & 3 \\ 2 & 2 \end{bmatrix}$ is $\lambda = -1$. What must be true about $A + I_2$ in order for -1 to be an eigenvalue of A ? How does this relate to $\det(A + I_2)$?
- (4) To find the eigenvalues of the matrix $A = \begin{bmatrix} 3 & 2 \\ 2 & 6 \end{bmatrix}$, we rewrite the equation $A\mathbf{x} = \lambda\mathbf{x}$ as $(A - \lambda I_2)\mathbf{x} = \mathbf{0}$. The coefficient matrix of this last system has the form $A - \lambda I_2 = \begin{bmatrix} 3 - \lambda & 2 \\ 2 & 6 - \lambda \end{bmatrix}$. The determinant of this matrix is a quadratic expression in λ . Since the eigenvalues will occur when the determinant is 0, we need to solve a quadratic equation. Find the resulting eigenvalues. (Note: One of the eigenvalues is 2.)
- (5) Can you explain why a 2×2 matrix can have at most two eigenvalues?



The Determinant of a Square Matrix

Around 1900 or so determinants were deemed much more important than they are today. In fact, determinants were used even before matrices. According to Tucker¹ determinants (not matrices) developed out of the study of coefficients of systems of linear equations and were used by Leibniz 150 years before the term matrix was coined by J. J. Sylvester in 1848. Even though determinants are not as important as they once were, the determinant of a matrix is still a useful quantity. We saw in Preview Activity 16.1 that the determinant of a matrix tells us if the matrix is invertible and how it can help us find eigenvalues. In this section, we will see how to find the determinant of any size matrix and how to use this determinant to find the eigenvalues.

The determinant of a 2×2 matrix $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ is $\det(A) = ad - bc$. The matrix A is invertible if and only if $\det(A) \neq 0$. We will use a recursive approach to find the determinants of larger size matrices building from the 2×2 determinants. We present the result in the 3×3 case here – a more detailed analysis can be found at the end of this section.

To find the determinant of a 3×3 matrix $A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$, we will use the determinants of three 2×2 matrices. More specifically, the determinant of A , denoted $\det(A)$ is the quantity

$$a_{11} \det \left(\begin{bmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{bmatrix} \right) - a_{12} \det \left(\begin{bmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{bmatrix} \right) + a_{13} \det \left(\begin{bmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{bmatrix} \right). \quad (16.1)$$

This sum is called a *cofactor expansion* of the determinant of A . The smaller matrices in this expansion are obtained by deleting certain rows and columns of the matrix A . In general, when finding the determinant of an $n \times n$ matrix, we find determinants of $(n - 1) \times (n - 1)$ matrices, which we can again reduce to smaller matrices to calculate.

We will use the specific matrix

$$A = \begin{bmatrix} 1 & 2 & 0 \\ 1 & 4 & 3 \\ 2 & 2 & 1 \end{bmatrix}$$

as an example in illustrating the cofactor expansion method in general.

- We first pick a row or column of A . We will pick the first row of A for this example.
- For each entry in the row (or column) we choose, in this case the first row, we will calculate the determinant of a smaller matrix obtained by removing the row and the column the entry is in. Let A_{ij} be the smaller matrix found by deleting the i th row and j th column of A . For entry a_{11} , we find the matrix A_{11} obtained by removing first row and first column:

$$A_{11} = \begin{bmatrix} 4 & 3 \\ 2 & 1 \end{bmatrix}.$$

For entry a_{12} , we find

$$A_{12} = \begin{bmatrix} 1 & 3 \\ 2 & 1 \end{bmatrix}.$$

¹Tucker, Alan. (1993). The Growing Importance of Linear Algebra in Undergraduate Mathematics. *The College Mathematics Journal*, 1, 3-9.



Finally, for entry a_{13} , we find

$$A_{13} = \begin{bmatrix} 1 & 4 \\ 2 & 2 \end{bmatrix}.$$

- Notice that in the 3×3 determinant formula in (16.1) above, the middle term had a (-) sign. The signs of the terms in the cofactor expansion alternate within each row and each column. More specifically, the sign of a term in the i th row and j th column is $(-1)^{i+j}$. We then obtain the following pattern of the signs within each row and column:

$$\begin{bmatrix} + & - & + & \cdots \\ - & + & - & \cdots \\ + & - & + & \cdots \\ \vdots & & & \end{bmatrix}$$

In particular, the sign factor for a_{11} is $(-1)^{1+1} = 1$, for a_{12} is $(-1)^{1+2} = -1$, and for a_{13} is $(-1)^{1+3} = 1$.

- For each entry a_{ij} in the row (or column) of A we chose, we multiply the entry a_{ij} by the determinant of A_{ij} and the sign $(-1)^{i+j}$. In this case, we obtain the following numbers

$$a_{11}(-1)^{1+1} \det(A_{11}) = 1 \det \begin{bmatrix} 4 & 3 \\ 2 & 1 \end{bmatrix} = 1(4 - 6) = -2$$

$$a_{12}(-1)^{1+2} \det(A_{12}) = -2 \det \begin{bmatrix} 1 & 3 \\ 2 & 1 \end{bmatrix} = -2(1 - 6) = 10$$

$$a_{13}(-1)^{1+3} \det(A_{13}) = 0$$

Note that in the last calculation, since $a_{13} = 0$, we did not have to evaluate the rest of the terms.

- Finally, we find the determinant by adding all these values:

$$\begin{aligned} \det(A) &= a_{11}(-1)^{1+1} \det(A_{11}) + a_{12}(-1)^{1+2} \det(A_{12}) \\ &\quad + a_{13}(-1)^{1+3} \det(A_{13}) \\ &= 8. \end{aligned}$$

Cofactors

We will now define the determinant of a general $n \times n$ matrix A in terms of a cofactor expansion as we did in the 3×3 case. To do so, we need some notation and terminology.

- We let A_{ij} be the submatrix of $A = [a_{ij}]$ found by deleting the i th row and j th column of A . The determinant of A_{ij} is called the ij th *minor* of A or the minor corresponding to the entry a_{ij} .



- Notice that in the 3×3 case, we used the opposite of the 1,2 minor in the sum. It will be the case that the terms in the cofactor expansion will alternate in sign. We can make the signs in the sum alternate by taking -1 to an appropriate power. As a result, we define the ij th *cofactor* C_{ij} of A as

$$C_{ij} = (-1)^{i+j} \det(A_{ij}).$$

- Finally, we define the determinant of A .

Definition 16.1. If $A = [a_{ij}]$ is an $n \times n$ matrix, the **determinant** of A is the scalar

$$\det(A) = a_{11}C_{11} + a_{12}C_{12} + a_{13}C_{13} + \cdots + a_{1n}C_{1n}$$

where $C_{ij} = (-1)^{i+j} \det(A_{ij})$ is the ij -cofactor of A and A_{ij} is the matrix obtained by removing row i and column j of matrix A .

This method for computing determinants is called the *cofactor expansion* or *Laplace expansion* of A along the 1st row. The cofactor expansion reduces the computation of the determinant of an $n \times n$ matrix to n computations of determinants of $(n - 1) \times (n - 1)$ matrices. These smaller matrices can be reduced again using cofactor expansions, so it can be a long and grueling process for large matrices. It turns out that we can actually take this expansion along any row or column of the matrix (a proof of this fact is given in Section 21). For example, the cofactor expansion along the 2nd row is

$$\det(A) = a_{21}C_{21} + a_{22}C_{22} + \cdots + a_{2n}C_{2n}$$

and along the 3rd column the formula is

$$\det(A) = a_{13}C_{13} + a_{23}C_{23} + \cdots + a_{n3}C_{n3}.$$

Note that when finding a cofactor expansion, choosing a row or column with many zeros makes calculations easier.

Activity 16.1.

- (a) Let $A = \begin{bmatrix} 1 & 2 & -1 \\ -2 & 0 & 4 \\ 6 & 3 & 0 \end{bmatrix}$. Use the cofactor expansion along the first row to calculate the determinant of A by hand.
- (b) Calculate $\det(A)$ by using a cofactor expansion along the second row where $A = \begin{bmatrix} 1 & 4 & 2 \\ 0 & 2 & 0 \\ 2 & 5 & 3 \end{bmatrix}$.
- (c) Calculate the determinant of $\begin{bmatrix} 1 & -2 & 3 \\ 0 & 4 & -3 \\ 0 & 0 & 8 \end{bmatrix}$.
- (d) Which determinant property can be used to calculate the determinant in part (c)? Explain how. (Determinant properties are included below for easy reference.)



- (e) Consider the matrix $A = \begin{bmatrix} 1 & 1 & 2 \\ 0 & 2 & 1 \\ 1 & 2 & 2 \end{bmatrix}$. Let B be the matrix which results when c times row 1 is added to row 2 of A . Evaluate the determinant of B by hand to check that it is equal to the determinant of A , which verifies one other determinant property (in a specific case).

As with any new idea, like the determinant, we must ask what properties are satisfied. We state the following theorem without proof for the time being. For the interested reader, the proof of many of these properties is given in Section 21 and others in the exercises.

Theorem 16.2. *Given $n \times n$ matrices A, B , the following hold:*

- (1) $\det(AB) = \det(A) \cdot \det(B)$, and in particular $\det(A^k) = (\det A)^k$ for any positive integer k .
- (2) $\det(A^T) = \det(A)$.
- (3) A is invertible if and only if $\det(A) \neq 0$.
- (4) If A is invertible, then $\det(A^{-1}) = (\det A)^{-1}$.
- (5) For a 2×2 matrix $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, $\det(A) = ad - bc$.
- (6) If A is upper/lower triangular, then $\det(A)$ is the product of the entries on the diagonal.
- (7) The determinant of a matrix is the product of the eigenvalues, with each eigenvalue repeated as many times as its multiplicity.
- (8) Effect of row operations:
 - Adding a multiple of a row to another does NOT change the determinant of the matrix.
 - Multiplying a row by a constant multiplies the determinant by the same constant.
 - Row swapping multiplies the determinant by (-1) .
- (9) If the row echelon form U of A is obtained by adding multiples of one row to another, and row swapping, then $\det(A)$ is equal to $\det(U)$ multiplied by $(-1)^r$ where r is the number of row swappings done during the row reduction.

Note that if we were to find the determinant of a 4×4 matrix using the cofactor method, we will calculate determinants of 4 matrices of size 3×3 , each of which will require 3 determinant calculations again. So, we will need a total of 12 calculations of determinants of 2×2 matrices. That is a lot of calculations. There are other, more efficient, methods for calculating determinants. For example, we can row reduce the matrix, keeping track of the effect that each row operation has on the determinant.



The Determinant of a 3×3 Matrix

Earlier we defined the determinant of a 3×3 matrix. In this section we endeavor to understand the motivation behind that definition.

We will repeat the process we went through in the 2×2 case to see how to define the determinant of a 3×3 matrix. Let

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}.$$

To find the inverse of A we augment A by the 3×3 identity matrix

$$[A | I_3] = \left[\begin{array}{ccc|cc|c} a_{11} & a_{12} & a_{13} & 1 & 0 & 0 \\ a_{21} & a_{22} & a_{23} & 0 & 1 & 0 \\ a_{31} & a_{32} & a_{33} & 0 & 0 & 1 \end{array} \right]$$

and row reduce the matrix (using appropriate technology) to obtain

$$\left[\begin{array}{ccc|ccc} 1 & 0 & 0 & \frac{a_{33}a_{22} - a_{32}a_{23}}{d} & -\frac{a_{33}a_{12} - a_{32}a_{13}}{d} & \frac{-a_{13}a_{22} + a_{12}a_{23}}{d} \\ 0 & 1 & 0 & -\frac{a_{33}a_{21} - a_{31}a_{23}}{d} & \frac{a_{33}a_{11} - a_{31}a_{13}}{d} & -\frac{a_{23}a_{11} - a_{21}a_{13}}{d} \\ 0 & 0 & 1 & \frac{-a_{31}a_{22} + a_{32}a_{21}}{d} & -\frac{a_{32}a_{11} - a_{31}a_{12}}{d} & \frac{a_{22}a_{11} - a_{21}a_{12}}{d} \end{array} \right],$$

where

$$d = a_{33}a_{11}a_{22} - a_{33}a_{21}a_{12} - a_{31}a_{13}a_{22} - a_{32}a_{11}a_{23} + a_{32}a_{21}a_{13} + a_{31}a_{12}a_{23}. \quad (16.2)$$

In this case, we can see that the inverse of the 3×3 matrix A will be defined if and only if $d \neq 0$. So, in the 3×3 case the determinant of A will be given by the value of d in Equation (16.2). What remains is for us to see how this is related to determinants of 2×2 sub-matrices of A .

To start, we collect all terms involving a_{11} in d . A little algebra shows that

$$\det(A) = a_{11}(a_{33}a_{22} - a_{32}a_{23}) - a_{33}a_{21}a_{12} - a_{31}a_{13}a_{22} + a_{32}a_{21}a_{13} + a_{31}a_{12}a_{23}.$$

Now let's collect the remaining terms involving a_{12} :

$$\det(A) = a_{11}(a_{33}a_{22} - a_{32}a_{23}) - a_{12}(a_{33}a_{21} - a_{31}a_{23}) - a_{31}a_{13}a_{22} + a_{32}a_{21}a_{13}.$$

Finally, we collect the terms involving a_{13} :

$$\det(A) = a_{11}(a_{33}a_{22} - a_{32}a_{23}) - a_{12}(a_{33}a_{21} - a_{31}a_{23}) + a_{13}(a_{32}a_{21} - a_{31}a_{22}).$$

Now we can connect the determinant of A to determinants of 2×2 sub-matrices of A .

- Notice that

$$a_{33}a_{22} - a_{32}a_{23}$$

is the determinant of the 2×2 matrix $\begin{bmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{bmatrix}$ obtained from A by deleting the first row and first column.



- Similarly, the expression

$$a_{33}a_{21} - a_{31}a_{23}$$

is the determinant of the 2×2 matrix $\begin{bmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{bmatrix}$ obtained from A by deleting the first row and second column.

- Finally, the expression

$$a_{32}a_{21} - a_{31}a_{22}$$

is the determinant of the 2×2 matrix $\begin{bmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{bmatrix}$ obtained from A by deleting the first row and third column.

Putting this all together gives us formula (16.1) for the determinant of a 3×3 matrix as we defined earlier.

Two Devices for Remembering Determinants

There are useful ways to remember how to calculate the formulas for determinants of 2×2 and 3×3 matrices. In the 2×2 case of $A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$, we saw that

$$|A| = a_{11}a_{22} - a_{21}a_{12}.$$

This makes $|A|$ the product of the diagonal elements a_{11} and a_{22} minus the product of the off-diagonal elements a_{12} and a_{21} . We can visualize this in an array by drawing arrows across the diagonal and off-diagonal, with a plus sign on the diagonal arrow indicating that we add the product of the diagonal elements and a minus sign on the off-diagonal arrow indicating that we subtract the product of the off-diagonal elements as shown in Figure 16.2.

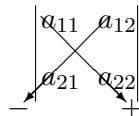


Figure 16.2: A diagram to remember the 2×2 determinant.

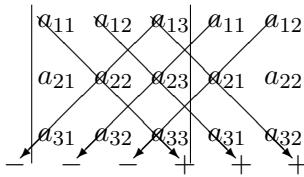
We can do a similar thing for the determinant of a 3×3 matrix. In this case, we extend the 3×3 array to a 3×5 array by adjoining the first two columns onto the matrix. We then add the products along the diagonals going from left to right and subtract the products along the diagonals going from right to left as indicated in Figure 16.3.

Examples

What follows are worked examples that use the concepts from this section.

Example 16.3. For each of the following



Figure 16.3: A diagram to remember the 3×3 determinant.

- Identify the sub-matrices $A_{1,j}$
- Determine the cofactors $C_{1,j}$.
- Use the cofactor expansion to calculate the determinant.

$$(a) A = \begin{bmatrix} 3 & 6 & 2 \\ 0 & 4 & -1 \\ 5 & 0 & 1 \end{bmatrix}$$

$$(b) A = \begin{bmatrix} 3 & 0 & 1 & 1 \\ 2 & 1 & 2 & 1 \\ 1 & -2 & 2 & -1 \\ -3 & 2 & 3 & 1 \end{bmatrix}$$

Example Solution.

- (a) With a 3×3 matrix, we will find the sub-matrices A_{11} , A_{12} , and A_{13} . Recall that A_{ij} is the sub-matrix of A obtained by deleting the i th row and j th column of A . Thus,

$$A_{11} = \begin{bmatrix} 4 & -1 \\ 0 & 1 \end{bmatrix} \quad A_{12} = \begin{bmatrix} 0 & -1 \\ 5 & 1 \end{bmatrix} \quad \text{and} \quad A_{13} = \begin{bmatrix} 0 & 4 \\ 5 & 0 \end{bmatrix}.$$

The ij th cofactor is $C_{ij} = (-1)^{i+j} \det(A_{ij})$, so

$$\begin{aligned} C_{11} &= (-1)^2 \begin{bmatrix} 4 & -1 \\ 0 & 1 \end{bmatrix} = 4 \\ C_{12} &= (-1)^3 \begin{bmatrix} 0 & -1 \\ 5 & 1 \end{bmatrix} = -5 \\ C_{13} &= (-1)^4 \begin{bmatrix} 0 & 4 \\ 5 & 0 \end{bmatrix} = -20. \end{aligned}$$

Then

$$\det(A) = a_{11}C_{11} + a_{12}C_{12} + a_{13}C_{13} = (3)(4) + (6)(-5) + (2)(-20) = -58.$$

(b) With a 4×4 matrix, we will find the sub-matrices A_{11} , A_{12} , A_{13} , and A_{14} . We see that

$$\begin{aligned} A_{11} &= \begin{bmatrix} 1 & 2 & 1 \\ -2 & 2 & -1 \\ 2 & 3 & 1 \end{bmatrix} \\ A_{12} &= \begin{bmatrix} 2 & 2 & 1 \\ 1 & 2 & -1 \\ -3 & 3 & 1 \end{bmatrix} \\ A_{13} &= \begin{bmatrix} 2 & 1 & 1 \\ 1 & -2 & -1 \\ -3 & 2 & 1 \end{bmatrix} \\ A_{14} &= \begin{bmatrix} 2 & 1 & 2 \\ 1 & -2 & 2 \\ -3 & 2 & 3 \end{bmatrix}. \end{aligned}$$

To calculate the ij th cofactor $C_{ij} = (-1)^{i+j} \det(A_{ij})$, we need to calculate the determinants of the A_{1j} . Using the device for calculating the determinant of a 3×3 matrix we have that

$$\begin{aligned} \det(A_{11}) &= \det \left(\begin{bmatrix} 1 & 2 & 1 \\ -2 & 2 & -1 \\ 2 & 3 & 1 \end{bmatrix} \right) \\ &= (1)(2)(1) + (2)(-1)(2) + (1)(-2)(3) \\ &\quad - (1)(2)(2) - (1)(-1)(3) - (2)(-2)(1) \\ &= -5, \end{aligned}$$

$$\begin{aligned} \det(A_{12}) &= \det \left(\begin{bmatrix} 2 & 2 & 1 \\ 1 & 2 & -1 \\ -3 & 3 & 1 \end{bmatrix} \right) \\ &= (2)(2)(1) + (2)(-1)(-3) + (1)(1)(3) \\ &\quad - (1)(2)(-3) - (2)(-1)(3) - (2)(1)(1) \\ &= 23, \end{aligned}$$

$$\begin{aligned} \det(A_{13}) &= \det \left(\begin{bmatrix} 2 & 1 & 1 \\ 1 & -2 & -1 \\ -3 & 2 & 1 \end{bmatrix} \right) \\ &= (2)(-2)(1) + (1)(-1)(-3) + (1)(1)(2) \\ &\quad - (1)(-2)(-3) - (2)(-1)(2) - (1)(1)(1) \\ &= -2, \end{aligned}$$

and

$$\begin{aligned}\det(A_{14}) &= \det \left(\begin{bmatrix} 2 & 1 & 2 \\ 1 & -2 & 2 \\ -3 & 2 & 3 \end{bmatrix} \right) \\ &= (2)(-2)(3) + (1)(2)(-3) + (2)(1)(2) \\ &\quad - (2)(-2)(-3) - (2)(2)(2) - (1)(1)(3) \\ &= -37.\end{aligned}$$

Then

$$\begin{aligned}C_{11} &= (-1)^2 \det(A_{11}) = -5 \\ C_{12} &= (-1)^3 \det(A_{12}) = -23 \\ C_{13} &= (-1)^4 \det(A_{13}) = -2 \\ C_{14} &= (-1)^5 \det(A_{14}) = 37\end{aligned}$$

and so

$$\begin{aligned}\det(B) &= b_{11}C_{11} + b_{12}C_{12} + b_{13}C_{13} + b_{14}C_{14} \\ &= (3)(-5) + (0)(-23) + (1)(-2) + (1)(37) \\ &= 20.\end{aligned}$$

Example 16.4. Show that for any 2×2 matrices A and B ,

$$\det(AB) = \det(A)\det(B).$$

Example Solution.

Let $A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$ and $B = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix}$. Then

$$AB = \begin{bmatrix} a_{11}b_{11} + a_{12}b_{21} & a_{11}b_{12} + a_{12}b_{22} \\ a_{21}b_{11} + a_{22}b_{21} & a_{21}b_{12} + a_{22}b_{22} \end{bmatrix}.$$

So

$$\begin{aligned}\det(AB) &= (a_{11}b_{11} + a_{12}b_{21})(a_{21}b_{12} + a_{22}b_{22}) \\ &\quad - (a_{11}b_{12} + a_{12}b_{22})(a_{21}b_{11} + a_{22}b_{21}) \\ &= (a_{11}b_{11}a_{21}b_{12} + a_{11}b_{11}a_{22}b_{22} + a_{12}b_{21}a_{21}b_{12} + a_{12}b_{21}a_{22}b_{22}) \\ &\quad - (a_{11}b_{12}a_{21}b_{11} + a_{11}b_{12}a_{22}b_{21} + a_{12}b_{22}a_{21}b_{11} + a_{12}b_{22}a_{22}b_{21}) \\ &= a_{11}b_{11}a_{22}b_{22} + a_{12}b_{21}a_{21}b_{12} - a_{11}b_{12}a_{22}b_{21} - a_{12}b_{22}a_{21}b_{11}.\end{aligned}$$

Also,

$$\begin{aligned}\det(A)\det(B) &= (a_{11}a_{22} - a_{12}a_{21})(b_{11}b_{22} - b_{12}b_{21}) \\ &= a_{11}a_{22}b_{11}b_{22} - a_{11}a_{22}b_{12}b_{21} - a_{12}a_{21}b_{11}b_{22} + a_{12}a_{21}b_{12}b_{21}.\end{aligned}$$

We conclude that $\det(AB) = \det(A)\det(B)$ if A and B are 2×2 matrices.



Summary

- The determinant of an $n \times n$ matrix $A = [a_{ij}]$ is found by taking the cofactor expansion of A along the first row. That is

$$\det(A) = a_{11}C_{11} + a_{12}C_{12} + a_{13}C_{13} + \cdots + a_{1n}C_{1n},$$

where

- A_{ij} is the sub-matrix of A found by deleting the i th row and j th column of A .
- $C_{ij} = (-1)^{i+j} \det(A_{ij})$ is the ij th *cofactor* of A .
- The matrix A is invertible if and only if $\det(A) \neq 0$.

Exercises

- (1) Use the cofactor expansion to explain why multiplying each of the entries of a 3×3 matrix A by 2 multiplies the determinant of A by 8.

- (2) Use the determinant criterion to determine for which c the matrix $A = \begin{bmatrix} 1 & 1 & 2 \\ 1 & 0 & c \\ 2 & -1 & 2 \end{bmatrix}$ is invertible.

- (3) Let A be a square matrix.

- (a) Explain why $\det(A^2) = [\det(A)]^2$

- (b) Expand on the argument from (a) to explain why $\det(A^k) = [\det(A)]^k$ for any positive integer k .

- (c) Suppose that A is an invertible matrix and k is a positive integer. Must A^k be an invertible matrix? Why or why not?

- (4) Let A be an invertible matrix. Explain why $\det(A^{-1}) = \frac{1}{\det(A)}$ using determinant properties.

- (5) Simplify the following determinant expression using determinant properties:

$$\det(PA^4P^{-1}A^T(A^{-1})^3)$$

- (6) Find the eigenvalues of the following matrices. Find a basis for and the dimension of each eigenspace.

(a) $A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$

(b) $A = \begin{bmatrix} 2 & 0 & 3 \\ 0 & 1 & 0 \\ 0 & 1 & 2 \end{bmatrix}$

(7) Label each of the following statements as True or False. Provide justification for your response.

- (a) **True/False** For any two $n \times n$ matrices A and B , $\det(A + B) = \det A + \det B$.
- (b) **True/False** For any square matrix A , $\det(-A) = -\det(A)$.
- (c) **True/False** For any square matrix A , $\det(-A) = \det(A)$.
- (d) **True/False** The determinant of a square matrix with all non-zero entries is non-zero.
- (e) **True/False** If the determinant of A is non-zero, then so is the determinant of A^2 .
- (f) **True/False** If the determinant of a matrix A is 0, then one of the rows of A is a linear combination of the other rows.
- (g) **True/False** For any square matrix A , $\det(A^2) > \det(A)$.
- (h) **True/False** If A and B are $n \times n$ matrices and AB is invertible, then A and B are invertible.
- (i) **True/False** If A^2 is the zero matrix, then the only eigenvalue of A is 0.
- (j) **True/False** If 0 is an eigenvalue of A , then 0 is an eigenvalue of AB for any B of the same size as A .
- (k) **True/False** Suppose A is a 3×3 matrix. Then any three eigenvectors of A will form a basis of \mathbb{R}^3 .

Project: Area and Volume Using Determinants

The approach we will take to connecting area (volume) to the determinant will help shed light on properties of the determinant that we will discuss from an algebraic perspective in a later section. First, we mention some basic properties of area (we focus on area for now, but these same properties are valid for volumes as well). As a shorthand, we denote the area of a region R by $\text{Area}(R)$.

- Area cannot be negative.
- If two regions R_1 and R_2 don't overlap, then the area of the union of the regions is equal to the sum of the areas of the regions. That is, if $R_1 \cap R_2 = \emptyset$, then $\text{Area}(R_1 \cup R_2) = \text{Area}(R_1) + \text{Area}(R_2)$.
- Area is invariant under translation. That is, if we move a geometric region by the same amount uniformly in a given direction, the area of the original region and the area of the transformed region are the same. A translation of a region is done by just adding a fixed vector to each vector in the region. That is, a translation by a vector \mathbf{v} is a function $T_{\mathbf{v}}$ such that the image $T_{\mathbf{v}}(R)$ of a region R is defined as

$$T_{\mathbf{v}}(R) = \{\mathbf{r} + \mathbf{v} : \mathbf{r} \in R\}.$$

Since area is translation invariant, $\text{Area}(T_{\mathbf{v}}(R)) = \text{Area}(R)$.



- The area of a one-dimensional object like a line segment is 0.

Now we turn our attention to areas of parallelograms. Let \mathbf{u} and \mathbf{v} be vectors in \mathbb{R}^2 . The parallelogram $P(\mathbf{u}, \mathbf{v})$ defined by \mathbf{u} and \mathbf{v} with point Q as basepoint is the set

$$P(\mathbf{u}, \mathbf{v}) = \{\overrightarrow{OQ} + r\mathbf{u} + s\mathbf{v} : 0 \leq r, s \leq 1\}.$$

An illustration of such a parallelogram is shown at left in Figure 16.4. If $\mathbf{u} = [u_1 \ u_2]^\top$ and $\mathbf{v} = [v_1 \ v_2]^\top$, then we will also represent $P(\mathbf{u}, \mathbf{v})$ as $P\left(\begin{bmatrix} u_1 & u_2 \\ v_1 & v_2 \end{bmatrix}\right)$.

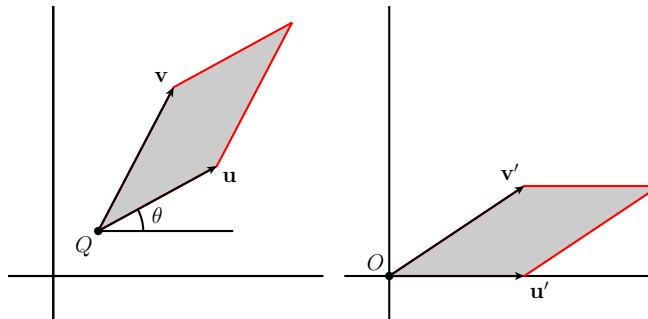


Figure 16.4: A parallelogram and a translated, rotated parallelogram.

$[v_1 \ v_2]^\top$, then we will also represent $P(\mathbf{u}, \mathbf{v})$ as $P\left(\begin{bmatrix} u_1 & u_2 \\ v_1 & v_2 \end{bmatrix}\right)$.

Since area is translation and rotation invariant, we can translate our parallelogram by $-\overrightarrow{OQ}$ to place its basepoint at the origin, then rotate by an angle θ (as shown at left in Figure 16.4). This transforms the vector \mathbf{v} to a vector \mathbf{v}' and the vector \mathbf{u} to a vector \mathbf{u}' as shown at right in Figure 16.4. With this in mind we can always assume that our parallelograms have one vertex at the origin, with \mathbf{u} along the x -axis, and \mathbf{v} in standard position. Now we can investigate how to calculate the area of a parallelogram.

Project Activity 16.1. There are two situations to consider when we want to find the area of a parallelogram determined by vectors \mathbf{u} and \mathbf{v} , both shown in Figure 16.5. The parallelogram will be determined by the lengths of these vectors.

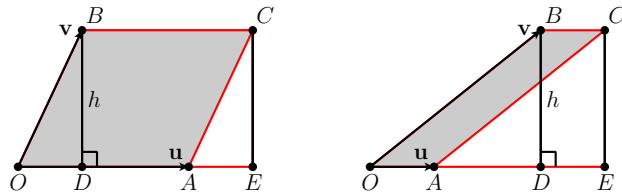


Figure 16.5: Parallelograms formed by \mathbf{u} and \mathbf{v}

- In the situation depicted at left in Figure 16.5, use geometry to explain why $\text{Area}(P(\mathbf{u}, \mathbf{v})) = h|\mathbf{u}|$. (Hint: What can we say about the triangles ODB and EAC ?)

- (b) In the situation depicted at right in Figure 16.5, use geometry to again explain why $\text{Area}(P(\mathbf{u}, \mathbf{v})) = h|\mathbf{u}|$. (Hint: What can we say about $\text{Area}(AEC)$ and $\text{Area}(ODB)$?)

The result of Project Activity 16.1 is that the area of $P(\mathbf{u}, \mathbf{v})$ is given by $h|\mathbf{u}|$, where h is the height of the parallelogram determined by dropping a perpendicular from the terminal point of \mathbf{v} to the line determined by the vector \mathbf{u} .

Now we turn to the question of how the determinant is related to area of a parallelogram. Our approach will use some properties of the area of $P(\mathbf{u}, \mathbf{v})$.

Project Activity 16.2. Let \mathbf{u} and \mathbf{v} be vectors that determine a parallelogram in \mathbb{R}^2 .

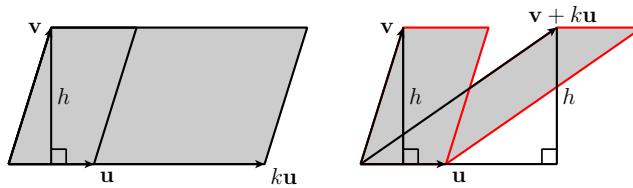


Figure 16.6: Parallelograms formed by $k\mathbf{u}$ and \mathbf{v} and by \mathbf{u} and $\mathbf{v} + k\mathbf{u}$.

- (a) Explain why

$$\text{Area}(P(\mathbf{u}, \mathbf{v})) = \text{Area}(P(\mathbf{v}, \mathbf{u})) \quad (16.3)$$

- (b) If k is any scalar, then $k\mathbf{u}$ either stretches or compresses \mathbf{u} . Use this idea, and the result of Project Activity 16.1, to explain why

$$\text{Area}(P(k\mathbf{u}, \mathbf{v})) = \text{Area}(P(\mathbf{u}, k\mathbf{v})) = |k|\text{Area}(P(\mathbf{u}, \mathbf{v})) \quad (16.4)$$

for any real number k . A representative picture of this situation is shown at left in Figure 16.5 for a value of $k > 1$. You will also need to consider what happens when $k < 0$.

- (c) Finally, use the result of Project Activity 16.1 to explain why

$$\text{Area}(P(\mathbf{u} + k\mathbf{v}, \mathbf{v})) = \text{Area}(P(\mathbf{u}, \mathbf{v} + k\mathbf{u})) = \text{Area}(P(\mathbf{u}, \mathbf{v})) \quad (16.5)$$

for any real number k . A representative picture is shown at right in Figure 16.6.

Properties (16.4) and (16.5) will allow us to calculate the area of the parallelogram determined by vectors \mathbf{u} and \mathbf{v} .

Project Activity 16.3. Let $\mathbf{u} = [u_1 \ u_2]^\top$ and $\mathbf{v} = [v_1 \ v_2]^\top$. We will now demonstrate that

$$\text{Area}(P(\mathbf{u}, \mathbf{v})) = \left| \det \begin{pmatrix} u_1 & u_2 \\ v_1 & v_2 \end{pmatrix} \right|.$$

Before we begin, note that if both u_1 and v_1 are 0, then \mathbf{u} and \mathbf{v} are parallel. This makes $P(\mathbf{u}, \mathbf{v})$ a line segment and so $\text{Area}(P(\mathbf{u}, \mathbf{v})) = 0$. But if $u_1 = v_1 = 0$, it is also the case that

$$\det \begin{pmatrix} u_1 & u_2 \\ v_1 & v_2 \end{pmatrix} = u_1 v_2 - u_2 v_1 = 0$$

as well. So we can assume that at least one of u_1, v_1 is not 0. Since $P(\mathbf{u}, \mathbf{v}) = P(\mathbf{v}, \mathbf{u})$, we can assume without loss of generality that $u_1 \neq 0$.

- (a) Explain using properties (16.4) and (16.5) as appropriate why

$$\text{Area}(P(\mathbf{u}, \mathbf{v})) = \text{Area} \left(P \left(\mathbf{u}, \begin{bmatrix} 0 & v_2 - \frac{v_1}{u_1} u_2 \end{bmatrix} \right) \right).$$

- (b) Let $\mathbf{v}_1 = \begin{bmatrix} 0 & v_2 - \frac{v_1}{u_1} u_2 \end{bmatrix}^\top$. Recall that our alternate representation of $P(\mathbf{u}, \mathbf{v})$) allows us to write

$$\text{Area}(P(\mathbf{u}, \mathbf{v}_1)) = \text{Area} \left(P \left(\begin{bmatrix} u_1 & u_2 \\ 0 & v_2 - \frac{v_1}{u_1} u_2 \end{bmatrix} \right) \right).$$

This should seem very suggestive. We are essentially applying the process of Gaussian elimination to our parallelogram matrix to reduce it to a diagonal matrix. From there, we can calculate the area. The matrix form should indicate the next step – applying an operation to eliminate the entry in the first row and second column. To do this, we need to consider what happens if $v_2 - \frac{v_1}{u_1} u_2 = 0$ and if $v_2 - \frac{v_1}{u_1} u_2 \neq 0$.

- i. Assume that $v_2 - \frac{v_1}{u_1} u_2 = 0$. Explain why $\text{Area}(P(\mathbf{u}, \mathbf{v})) = 0$. Then explain why

$$\text{Area}(P(\mathbf{u}, \mathbf{v})) = 0 = \det \left(\begin{bmatrix} u_1 & u_2 \\ v_1 & v_2 \end{bmatrix} \right).$$

- ii. Now we consider the case when $v_2 - \frac{v_1}{u_1} u_2 \neq 0$. Complete the process as in part (a), using properties (16.4) and (16.5) (compare to Gaussian elimination) to continue to reduce the problem of calculating $\text{Area}(P(\mathbf{u}, \mathbf{v}))$ to one of calculating $\text{Area}(P(\mathbf{e}_1, \mathbf{e}_2))$. Use this process to conclude that

$$\text{Area}(P(\mathbf{u}, \mathbf{v})) = \left| \det \left(\begin{bmatrix} u_1 & u_2 \\ v_1 & v_2 \end{bmatrix} \right) \right|.$$

We can apply the same arguments as above using rotations, translations, shearings, and scalings to show that the properties of area given above work in any dimension. Given vectors $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n$ in \mathbb{R}^n , we let

$$P(\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n) = \{\overrightarrow{OQ} + x_1 \mathbf{u}_1 + x_2 \mathbf{u}_2 + \dots + x_n \mathbf{u}_n : 0 \leq x_i \leq 1 \text{ for each } i\}.$$

If $n = 2$, then $P(\mathbf{u}_1, \mathbf{u}_2)$ is the parallelogram determined by \mathbf{u}_1 and \mathbf{u}_2 with basepoint Q . If $n = 3$, then $P(\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3)$ is the parallelepiped with basepoint Q determined by $\mathbf{u}_1, \mathbf{u}_2$, and \mathbf{u}_3 . In higher dimensions the sets $P(\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n)$ are called parallelotopes, and we use the notation $\text{Vol}(P(\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n))$ for their volume. The n -dimensional volumes of these parallelotopes satisfy the following properties:

$$\begin{aligned} \text{Vol}(P(\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_{i-1}, \mathbf{u}_i, \mathbf{u}_{i+1}, \dots, \mathbf{u}_{j-1}, \mathbf{u}_j, \mathbf{u}_{j+1}, \dots, \mathbf{u}_n)) \\ = \text{Vol}(P(\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_{i-1}, \mathbf{u}_j, \mathbf{u}_{i+1}, \dots, \mathbf{u}_{j-1}, \mathbf{u}_i, \mathbf{u}_{j+1}, \dots, \mathbf{u}_n)) \end{aligned} \quad (16.6)$$

for any i and j .

$$\text{Vol}(P(\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_{i-1}, k\mathbf{u}_i, \mathbf{u}_{i+1}, \dots, \mathbf{u}_n)) = |k| \text{Vol}(P(\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n)) \quad (16.7)$$

for any real number k and any i .

$$\text{Vol}(P(\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_{i-1}, \mathbf{u}_i + k\mathbf{u}_j, \mathbf{u}_{i+1}, \dots, \mathbf{u}_n)) = \text{Vol}(P(\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n)) \quad (16.8)$$

for any real number k and any distinct i and j .



Project Activity 16.4. We now show that $\text{Vol}(P(\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3))$ is the absolute value of the determinant of $\begin{bmatrix} \mathbf{u}_1 \\ \mathbf{u}_2 \\ \mathbf{u}_3 \end{bmatrix}$. For easier notation, let $\mathbf{u} = [u_1 \ u_2 \ u_3]^\top$, $\mathbf{v} = [v_1 \ v_2 \ v_3]^\top$, and $\mathbf{w} = [w_1 \ w_2 \ w_3]^\top$.

As we argued in the 2-dimensional case, we can assume that all terms that we need to be nonzero are nonzero, and we can do so without verification.

- (a) Explain how property (16.7) shows that $\text{Vol}(P(\mathbf{u}, \mathbf{v}, \mathbf{w}))$ is equal to

$$\text{Vol} \left(P \left(\begin{bmatrix} u_1 & u_2 & u_3 \\ 0 & \frac{1}{u_1}(v_2u_1 - v_1u_2) & \frac{1}{u_1}(v_3u_1 - v_1u_3) \\ 0 & \frac{1}{u_1}(w_2u_1 - w_1u_2) & \frac{1}{u_1}(w_3u_1 - w_1u_3) \end{bmatrix} \right) \right).$$

(Hint: Think about how these properties are related to row operations.)

- (b) Now let $\mathbf{v}_1 = \left[0 \ \frac{1}{u_1}(v_2u_1 - v_1u_2) \ \frac{1}{u_1}(v_3u_1 - v_1u_3) \right]^\top$ and $\mathbf{w}_1 = \left[0 \ \frac{1}{u_1}(w_2u_1 - w_1u_2) \ \frac{1}{u_1}(w_3u_1 - w_1u_3) \right]^\top$. Explain how property (16.7) shows that $\text{Vol}(P(\mathbf{u}, \mathbf{v}, \mathbf{w}))$ is equal to

$$\text{Vol} \left(P \left(\begin{bmatrix} u_1 & u_2 & u_3 \\ 0 & \frac{1}{u_1}(v_2u_1 - v_1u_2) & \frac{1}{u_1}(v_3u_1 - v_1u_3) \\ 0 & 0 & d \end{bmatrix} \right) \right),$$

where

$$d = \frac{1}{u_1v_2 - u_2v_1}(u_1(v_2w_3 - v_3w_2) - u_2(v_1w_3 - v_3w_1) + u_3(v_1w_2 - v_2w_1)).$$

- (c) Just as we saw in the 2-dimensional case, we can proceed to use the diagonal entries to eliminate the entries above the diagonal without changing the volume to see that

$$\text{Vol}(P(\mathbf{u}, \mathbf{v}, \mathbf{w})) = \text{Vol} \left(P \left(\begin{bmatrix} u_1 & 0 & 0 \\ 0 & \frac{1}{u_1}(v_2u_1 - v_1u_2) & 0 \\ 0 & 0 & d \end{bmatrix} \right) \right).$$

Complete the process, applying appropriate properties to explain why

$$\text{Vol}(P(\mathbf{u}, \mathbf{v}, \mathbf{w})) = x \text{Vol}(P(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3))$$

for some constant x . Find the constant and, as a result, find a specific expression for $\text{Vol}(P(\mathbf{u}, \mathbf{v}, \mathbf{w}))$ involving a determinant.

Properties (16.6), (16.7), and (16.8) involve the analogs of row operations on matrices, and we will prove algebraically that the determinant exhibits the same properties. In fact, the determinant can be uniquely defined by these properties. So in a sense, the determinant is an area or volume function.

Section 17

The Characteristic Equation

Focus Questions

By the end of this section, you should be able to give precise and thorough answers to the questions listed below. You may want to keep these questions in mind to focus your thoughts as you complete the section.

- What is the characteristic polynomial of a matrix?
- What is the characteristic equation of a matrix?
- How and why is the characteristic equation of a matrix useful?
- How many different eigenvalues can an $n \times n$ matrix have?
- How large can the dimension of the eigenspace corresponding to an eigenvalue be?

Application: Modeling the Second Law of Thermodynamics

Pour cream into your cup of coffee and the cream spreads out; straighten up your room and it soon becomes messy again; when gasoline is mixed with air in a car's cylinders, it explodes if a spark is introduced. In each of these cases a transition from a low energy state (your room is straightened up) to a higher energy state (a messy, disorganized room) occurs. This can be described by entropy – a measure of the energy in a system. Low energy is organized (like ice cubes) and higher energy is not (like water vapor). It is a fundamental property of energy (as described by the second law of thermodynamics) that the entropy of a system cannot decrease. In other words, in the absence of any external intervention, things never become more organized.

The Ehrenfest model¹ is a Markov process proposed to explain the statistical interpretation of the second law of thermodynamics using the diffusion of gas molecules. This process can be modeled as a problem of balls and bins, as we will do later in this section. The characteristic

¹named after Paul and Tatiana Ehrenfest who introduced it in “Über zwei bekannte Einwände gegen das Boltzmannsche H-Theorem,” *Physikalische Zeitschrift*, vol. 8 (1907), pp. 311-314)

polynomial of the transition matrix will help us find the eigenvalues and allow us to analyze our model.

Introduction

We have seen that the eigenvalues of an $n \times n$ matrix A are the scalars λ so that $A - \lambda I_n$ has a nontrivial null space. Since a matrix has a nontrivial null space if and only if the matrix is not invertible, we can also say that λ is an eigenvalue of A if

$$\det(A - \lambda I_n) = 0. \quad (17.1)$$

This equation is called the *characteristic equation* of A . It provides us an *algebraic* way to find eigenvalues, which can then be used in finding eigenvectors corresponding to each eigenvalue.

Suppose we want to find the eigenvalues of $A = \begin{bmatrix} 1 & 1 \\ 1 & 3 \end{bmatrix}$. Note that

$$A - \lambda I_2 = \begin{bmatrix} 1 - \lambda & 1 \\ 1 & 3 - \lambda \end{bmatrix},$$

with determinant $(1 - \lambda)(3 - \lambda) - 1 = \lambda^2 - 4\lambda + 2$. Hence, the eigenvalues λ_1, λ_2 are the solutions of the characteristic equation $\lambda^2 - 4\lambda + 2 = 0$. Using quadratic formula, we find that $\lambda_1 = 2 + \sqrt{2}$ and $\lambda_2 = 2 - \sqrt{2}$ are the eigenvalues.

In this activity, our goal will be to use the characteristic equation to obtain information about eigenvalues and eigenvectors of a matrix with real entries.

Preview Activity 17.1.

- (1) For each of the following parts, use the characteristic equation to determine the eigenvalues of A . Then, for each eigenvalue λ , find a basis of the corresponding eigenspace, i.e., $\text{Nul}(A - \lambda I)$. You might want to recall how to find a basis for the null space of a matrix from Section 13. Also, make sure that your eigenvalue candidate λ yields nonzero eigenvectors in $\text{Nul}(A - \lambda I)$ for otherwise λ will not be an eigenvalue.

$$(a) A = \begin{bmatrix} 2 & 0 \\ 0 & -3 \end{bmatrix} \quad (b) A = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix} \quad (c) A = \begin{bmatrix} 1 & 4 \\ 2 & 3 \end{bmatrix}$$

- (2) Use your eigenvalue and eigenvector calculations of the above problem as a guidance to answer the following questions about a matrix with real entries.

- (a) At most how many eigenvalues can a 2×2 matrix have? Is it possible to have no eigenvalues? Is it possible to have only one eigenvalue? Explain.
- (b) If a matrix is an upper-triangular matrix (i.e., all entries below the diagonal are 0's, as in the first two matrices of the previous problem), what can you say about its eigenvalues? Explain.
- (c) How many linearly independent eigenvectors can be found for a 2×2 matrix? Is it possible to have a matrix without 2 linearly independent eigenvectors? Explain.

- (3) Using the characteristic equation, determine which matrices have 0 as an eigenvalue.



The Characteristic Equation

Until now, we have been given eigenvalues or eigenvectors of a matrix and determined eigenvectors and eigenvalues from the known information. In this section we use determinants to find (or approximate) the eigenvalues of a matrix. From there we can find (or approximate) the corresponding eigenvectors. The tool we will use is a polynomial equation, the *characteristic equation*, of a square matrix whose roots are the eigenvalues of the matrix. The characteristic equation will then provide us with an *algebraic* way of finding the eigenvalues of a square matrix.

We have seen that the eigenvalues of a square matrix A are the scalars λ so that $A - \lambda I$ has a nontrivial null space. Since a matrix has a nontrivial null space if and only if the matrix is not invertible, we can also say that λ is an eigenvalue of A if

$$\det(A - \lambda I) = 0. \quad (17.2)$$

Note that if A is an $n \times n$ matrix, then $\det(A - \lambda I)$ is a polynomial of degree n . Furthermore, if A has real entries, the polynomial has real coefficients. This polynomial, and the equation (17.2) are given special names.

Definition 17.1. Let A be an $n \times n$ matrix. The **characteristic polynomial** of A is the polynomial

$$\det(A - \lambda I_n),$$

where I_n is the $n \times n$ identity matrix. The **characteristic equation** of A is the equation

$$\det(A - \lambda I_n) = 0.$$

So the characteristic equation of A gives us an algebraic way of finding the eigenvalues of A .

Activity 17.1.

- (a) Find the characteristic polynomial of the matrix $A = \begin{bmatrix} 3 & -2 & 5 \\ 1 & 0 & 7 \\ 0 & 0 & 1 \end{bmatrix}$, and use the characteristic polynomial to find all of the eigenvalues of A .

- (b) Verify that 1 and 2 are the only eigenvalues of the matrix $\begin{bmatrix} 1 & 0 & 0 & 1 \\ 1 & 2 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$.

As we argued in Preview Activity 17.1, a 2×2 matrix can have at most 2 eigenvalues. For an $n \times n$ matrix, the characteristic polynomial will be a degree n polynomial, and we know from algebra that a degree n polynomial can have at most n roots. Since an eigenvalue of a matrix is a root of the characteristic polynomial of that matrix, we can conclude that an $n \times n$ matrix can have at most n distinct eigenvalues. Activity 17.1 (b) shows that a 4×4 matrix may have fewer than 4 eigenvalues, however. Note that one of these eigenvalues, the eigenvalue 1, appears three times as a root of the characteristic polynomial of the matrix. The number of times an eigenvalue appears as a root of the characteristic polynomial is called the (*algebraic*) *multiplicity* of the eigenvalue. More formally:



Definition 17.2. The **(algebraic) multiplicity** of an eigenvalue λ of a matrix A is the largest integer m so that $(x - \lambda)^m$ divides the characteristic polynomial of A .

Thus, in Activity 17.1 (b) the eigenvalue 1 has multiplicity 3 and the eigenvalue 2 has multiplicity 1. Notice that if we count the eigenvalues of an $n \times n$ matrix with their multiplicities, the total will always be n .

If A is a matrix with real entries, then the characteristic polynomial will have real coefficients. It is possible that the characteristic polynomial can have complex roots, and that the matrix A has complex eigenvalues. The Fundamental Theorem of Algebra shows us that if a real matrix has complex eigenvalues, then those eigenvalues will appear in conjugate pairs, i.e., if $\lambda_1 = a + ib$ is an eigenvalue of A , then $\lambda_2 = a - ib$ is another eigenvalue of A . Furthermore, for an odd degree polynomial, since the complex eigenvalues will come in conjugate pairs, we will be able to find at least one real eigenvalue.

We now summarize the information we have so far about eigenvalues of an $n \times n$ real matrix:

Theorem 17.3. Let A be an $n \times n$ matrix with real entries. Then

- (1) There are at most n eigenvalues of A . If each eigenvalue (including complex eigenvalues) is counted with its multiplicity, there are exactly n eigenvalues.
- (2) If A has a complex eigenvalue λ , the complex conjugate of λ is also an eigenvalue of A .
- (3) If n is odd, A has at least one real eigenvalue.
- (4) If A is upper or lower-triangular, the eigenvalues are the entries on the diagonal.

Eigenspaces, A Geometric Example

Recall that for each eigenvalue λ of an $n \times n$ matrix A , the eigenspace of A corresponding to the eigenvalue λ is $\text{Nul}(A - \lambda I_n)$. These eigenspaces can tell us important information about the matrix transformation defined by A . For example, consider the matrix transformation T from \mathbb{R}^3 to \mathbb{R}^3 defined by $T(\mathbf{x}) = A\mathbf{x}$, where

$$A = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 2 \end{bmatrix}.$$

We are interested in understanding what this matrix transformation does to vectors in \mathbb{R}^3 . First we note that A has eigenvalues $\lambda_1 = 1$ and $\lambda_2 = 2$, with λ_1 having multiplicity 2. There is a pair

$\mathbf{v}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$ and $\mathbf{v}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$ of linearly independent eigenvectors for A corresponding to the

eigenvalue λ_1 and an eigenvector $\mathbf{v}_3 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ for A corresponding to the eigenvalue λ_2 . Note

that the vectors \mathbf{v}_1 , \mathbf{v}_2 , and \mathbf{v}_3 are linearly independent (recall from Theorem that eigenvectors corresponding to different eigenvalues are always linearly independent). So any vector \mathbf{b} in \mathbb{R}^3 can



be written uniquely as a linear combination of \mathbf{v}_1 , \mathbf{v}_2 , and \mathbf{v}_3 . Let's now consider the action of the matrix transformation T on a linear combination of \mathbf{v}_1 , \mathbf{v}_2 , and \mathbf{v}_3 . Note that

$$\begin{aligned} T(c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + c_3\mathbf{v}_3) &= c_1T(\mathbf{v}_1) + c_2T(\mathbf{v}_2) + c_3T(\mathbf{v}_3) \\ &= c_1\lambda_1\mathbf{v}_1 + c_2\lambda_1\mathbf{v}_2 + c_3\lambda_2\mathbf{v}_3 \\ &= (1)(c_1\mathbf{v}_1 + c_2\mathbf{v}_2) + (2)c_3\mathbf{v}_3. \end{aligned} \quad (17.3)$$

Equation (17.3) illustrates that it is most convenient to view the action of T in the coordinate system where $\text{Span}\{\mathbf{v}_1\}$ serves as the x -axis, $\text{Span}\{\mathbf{v}_2\}$ serves as the y -axis, and $\text{Span}\{\mathbf{v}_3\}$ as the z -axis. In this case, we can visualize that when we apply the transformation T to a vector $\mathbf{b} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + c_3\mathbf{v}_3$ in \mathbb{R}^3 the result is an output vector that is unchanged in the \mathbf{v}_1 - \mathbf{v}_2 plane and scaled by a factor of 2 in the \mathbf{v}_3 direction. For example, consider the box whose sides are determined by the vectors \mathbf{v}_1 , \mathbf{v}_2 , and \mathbf{v}_3 as shown in Figure 17.1. The transformation T stretches this box by a factor of 2 in the \mathbf{v}_3 direction and leaves everything else alone, as illustrated in Figure 17.1. So the entire $\text{Span}\{\mathbf{v}_1, \mathbf{v}_2\}$ is unchanged by T , but $\text{Span}\{\mathbf{v}_3\}$ is scaled by 2. In this situation, the eigenvalues and eigenvectors provide the most convenient perspective through which to visualize the action of the transformation T .

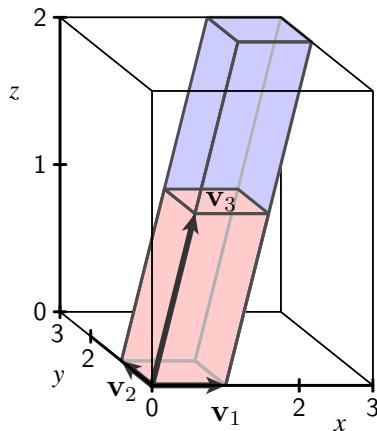


Figure 17.1: A box and a transformed box.

This geometric perspective illustrates how each eigenvalue and the corresponding eigenspace of A tells us something important about A . So it behooves us to learn a little more about eigenspaces.

Dimensions of Eigenspaces

There is a connection between the dimension of the eigenspace of a matrix corresponding to an eigenvalue and the multiplicity of that eigenvalue as a root of the characteristic polynomial. Recall that the dimension of a subspace of \mathbb{R}^n is the number of vectors in a basis for the eigenspace. We investigate the connection between dimension and multiplicity in the next activity.

Activity 17.2.

- (a) Find the dimension of the eigenspace for each eigenvalue of matrix $A = \begin{bmatrix} 3 & -2 & 5 \\ 1 & 0 & 7 \\ 0 & 0 & 1 \end{bmatrix}$ from Activity 17.1 (a).

- (b) Find the dimension of the eigenspace for each eigenvalue of matrix $A = \begin{bmatrix} 1 & 0 & 0 & 1 \\ 1 & 2 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$ from Activity 17.1 (b).

- (c) Consider now a 3×3 matrix with 3 distinct eigenvalues $\lambda_1, \lambda_2, \lambda_3$.
- Recall that a polynomial of degree can have at most three distinct roots. What does that say about the multiplicities of $\lambda_1, \lambda_2, \lambda_3$?
 - Use the fact that eigenvectors corresponding to distinct eigenvalues are linearly independent to find the dimensions of the eigenspaces for $\lambda_1, \lambda_2, \lambda_3$.

The examples in Activity 17.2 all provide instances of the principle that the dimension of an eigenspace corresponding to an eigenvalue λ cannot exceed the multiplicity of λ . Specifically:

Theorem 17.4. *If λ is an eigenvalue of A , the dimension of the eigenspace corresponding to λ is less than or equal to the multiplicity of λ .*

The examples we have seen raise another important point. The matrix $A = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 2 \end{bmatrix}$ from

our geometric example has two eigenvalues 1 and 2, with the eigenvalue 1 having multiplicity 2. If we let E_λ represent the eigenspace of A corresponding to the eigenvalue λ , then $\dim(E_1) = 2$

and $\dim(E_2) = 1$. If we change this matrix slightly to the matrix $B = \begin{bmatrix} 2 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}$ we see that

B has two eigenvalues 1 and 2, with the eigenvalue 1 having multiplicity 2. However, in this case we have $\dim(E_1) = 1$ (like the example in from Activities 17.1 (a) and 17.2 (a)). In this case the vector $\mathbf{v}_1 = [1 \ 0 \ 0]^\top$ forms a basis for E_2 and the vector $\mathbf{v}_2 = [0 \ 1 \ 0]^\top$ forms a basis for E_1 . We can visualize the action of B on the square formed by \mathbf{v}_1 and \mathbf{v}_2 in the xy -plane as a scaling by 2 in the \mathbf{v}_1 direction as shown in Figure 17.2, but since we do not have a third linearly independent eigenvector, the action of B in the direction of $[0 \ 0 \ 1]^\top$ is not so clear.

So the action of a matrix transformation can be more easily visualized if the dimension of each eigenspace is equal to the multiplicity of the corresponding eigenvalue. This geometric perspective leads us to define the geometric multiplicity of an eigenvalue.

Definition 17.5. The **geometric multiplicity** of an eigenvalue of an $n \times n$ matrix A is the dimension of the corresponding eigenspace $\text{Nul}(A - \lambda I_n)$.

Examples

What follows are worked examples that use the concepts from this section.



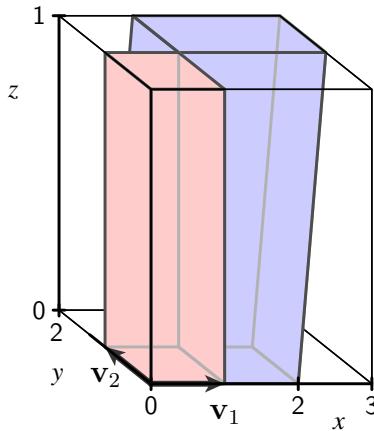


Figure 17.2: A box and a transformed box.

Example 17.6. Let $A = \begin{bmatrix} -1 & 0 & -2 \\ 2 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix}$.

- (a) Find the characteristic polynomial of A .
- (b) Factor the characteristic polynomial and find the eigenvalues of A .
- (c) Find a basis for each eigenspace of A .
- (d) Is it possible to find a basis for \mathbb{R}^3 consisting of eigenvectors of A ? Explain.

Example Solution.

- (a) The characteristic polynomial of A is

$$\begin{aligned} p(\lambda) &= \det(A - \lambda I_3) \\ &= \det \left(\begin{bmatrix} -1 - \lambda & 0 & -2 \\ 2 & 1 - \lambda & 2 \\ 0 & 0 & 1 - \lambda \end{bmatrix} \right) \\ &= (-1 - \lambda)(1 - \lambda)(1 - \lambda). \end{aligned}$$

- (b) The eigenvalues of A are the solutions to the characteristic equation. Since

$$p(\lambda) = (-1 - \lambda)(1 - \lambda)(1 - \lambda) = 0$$

implies $\lambda = -1$ or $\lambda = 1$, the eigenvalues of A are 1 and -1 .

- (c) To find a basis for the eigenspace of A corresponding to the eigenvalue 1, we find a basis for $\text{Nul}(A - I_3)$. The reduced row echelon form of $A - I_3 = \begin{bmatrix} -2 & 0 & -2 \\ 2 & 0 & 2 \\ 0 & 0 & 0 \end{bmatrix}$ is

$\begin{bmatrix} 1 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$. If $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$, then $(A - I_3)\mathbf{x} = \mathbf{0}$ has general solution

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -x_3 \\ x_2 \\ x_3 \end{bmatrix} = x_2 \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}.$$

Therefore, $\{[0 \ 1 \ 0]^\top, [-1 \ 0 \ 1]^\top\}$ is a basis for the eigenspace of A corresponding to the eigenvalue 1.

To find a basis for the eigenspace of A corresponding to the eigenvalue -1 , we find a basis for $\text{Nul}(A + I_3)$. The reduced row echelon form of $A + I_3 = \begin{bmatrix} 0 & 0 & -2 \\ 2 & 2 & 2 \\ 0 & 0 & 2 \end{bmatrix}$ is

$\begin{bmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$. If $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$, then $(A + I_3)\mathbf{x} = \mathbf{0}$ has general solution

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -x_2 \\ x_2 \\ 0 \end{bmatrix} = x_2 \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}.$$

Therefore, a basis for the eigenspace of A corresponding to the eigenvalue -1 is $\{[-1 \ 1 \ 0]^\top\}$.

- (d) Let $\mathbf{v}_1 = [0 \ 1 \ 0]^\top, [-1 \ 0 \ 1]^\top, \mathbf{v}_2 = [-1 \ 0 \ 1]^\top$, and $\mathbf{v}_3 = [-1 \ 1 \ 0]^\top$. Since eigenvectors corresponding to different eigenvalues are linearly independent, and since neither \mathbf{v}_1 nor \mathbf{v}_2 is a scalar multiple of the other, we can conclude that the set $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is a linearly independent set with $3 = \dim(\mathbb{R}^3)$ vectors. Therefore, $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is a basis for \mathbb{R}^3 consisting of eigenvectors of A .

Example 17.7. Find a 3×3 matrix A that has an eigenvector $\mathbf{v}_1 = [1 \ 0 \ 1]^\top$ with corresponding eigenvalue $\lambda_1 = 2$, an eigenvector $\mathbf{v}_2 = [0 \ 2 \ -3]^\top$ with corresponding eigenvalue $\lambda_2 = -3$, and an eigenvector $\mathbf{v}_3 = [-4 \ 0 \ 5]^\top$ with corresponding eigenvalue $\lambda_3 = 5$. Explain your process.

Example Solution. We are looking for a 3×3 matrix A such that $A\mathbf{v}_1 = 2\mathbf{v}_1, A\mathbf{v}_2 = -3\mathbf{v}_2$ and $A\mathbf{v}_3 = 5\mathbf{v}_3$. Since $\mathbf{v}_1, \mathbf{v}_2$, and \mathbf{v}_3 are eigenvectors corresponding to different eigenvalues, $\mathbf{v}_1, \mathbf{v}_2$,



and \mathbf{v}_3 are linearly independent. So the matrix $[\mathbf{v}_1 \mathbf{v}_2 \mathbf{v}_3]$ is invertible. It follows that

$$\begin{aligned} A[\mathbf{v}_1 \mathbf{v}_2 \mathbf{v}_3] &= [A\mathbf{v}_1 A\mathbf{v}_2 A\mathbf{v}_3] \\ A \begin{bmatrix} 1 & 0 & -4 \\ 0 & 2 & 0 \\ 1 & -3 & 5 \end{bmatrix} &= [2\mathbf{v}_1 - 3\mathbf{v}_2 5\mathbf{v}_3] \\ A \begin{bmatrix} 1 & 0 & -4 \\ 0 & 2 & 0 \\ 1 & -3 & 5 \end{bmatrix} &= \begin{bmatrix} 2 & 0 & -20 \\ 0 & -6 & 0 \\ 2 & 9 & 25 \end{bmatrix} \\ A = \begin{bmatrix} 2 & 0 & -20 \\ 0 & -6 & 0 \\ 2 & 9 & 25 \end{bmatrix} &\left[\begin{bmatrix} 1 & 0 & -4 \\ 0 & 2 & 0 \\ 1 & -3 & 5 \end{bmatrix}^{-1} \right] \\ A = \begin{bmatrix} 2 & 0 & -20 \\ 0 & -6 & 0 \\ 2 & 9 & 25 \end{bmatrix} &\left[\begin{bmatrix} \frac{5}{9} & \frac{2}{3} & \frac{4}{9} \\ 0 & \frac{1}{2} & 0 \\ -\frac{1}{9} & \frac{1}{6} & \frac{1}{9} \end{bmatrix} \right] \\ A = \begin{bmatrix} \frac{10}{3} & -2 & -\frac{4}{3} \\ 0 & -3 & 0 \\ -\frac{5}{3} & 10 & \frac{11}{3} \end{bmatrix}. & \end{aligned}$$

Summary

In this section we studied the characteristic polynomial of a matrix and similar matrices.

- If A is an $n \times n$ matrix, the characteristic polynomial of A is the polynomial

$$\det(A - \lambda I_n),$$

where I_n is the $n \times n$ identity matrix.

- If A is an $n \times n$ matrix, the characteristic equation of A is the equation

$$\det(A - \lambda I_n) = 0.$$

- The characteristic equation of a square matrix provides us an algebraic method to find the eigenvalues of the matrix.
- The eigenvalues of an upper or lower-triangular matrix are the entries on the diagonal.
- There are at most n eigenvalues of an $n \times n$ matrix.
- For a real matrix A , if an eigenvalue λ of A is complex, then the complex conjugate of λ is also an eigenvalue.
- The algebraic multiplicity of an eigenvalue λ is the multiplicity of λ as a root of the characteristic equation.
- The dimension of the eigenspace corresponding to an eigenvalue λ is less than or equal to the algebraic multiplicity of λ .



Exercises

- (1) There is interesting relationship² between a matrix and its characteristic equation that we explore in this exercise.

- (a) We first illustrate with an example. Let $B = \begin{bmatrix} 1 & 2 \\ 1 & -2 \end{bmatrix}$.
- Show that $\lambda^2 + \lambda - 4$ is the characteristic polynomial for B .
 - Calculate B^2 . Then compute $B^2 + B - 4I_2$. What do you get?
- (b) The first part of this exercise presents an example of a matrix that satisfies its own characteristic equation. Explain for a general $n \times n$ matrix, why A satisfies its characteristic equation.
- (2) There is a useful relationship between the determinant and eigenvalues of a matrix A that we explore in this exercise.
- (a) Let $B = \begin{bmatrix} 2 & 3 \\ 8 & 4 \end{bmatrix}$. Find the determinant of B and the eigenvalues of B , and compare $\det(B)$ to the eigenvalues of B .
- (b) Let A be an $n \times n$ matrix. In this part of the exercise we argue the general case illustrated in the previous part – that $\det(A)$ is the product of the eigenvalues of A . Let $p(\lambda) = \det(A - \lambda I_n)$ be the characteristic polynomial of A .
- Let $\lambda_1, \lambda_2, \dots, \lambda_n$ be the eigenvalues of A (note that these eigenvalues may not all be distinct). Recall that if r is a root of a polynomial $q(x)$, then $(x - r)$ is a factor of $q(x)$. Use this idea to explain why

$$p(\lambda) = (-1)^n(\lambda - \lambda_1)(\lambda - \lambda_2) \cdots (\lambda - \lambda_n).$$

- Explain why $p(0) = \lambda_1 \lambda_2 \cdots \lambda_n$.
- Why is $p(0)$ also equal to $\det(A)$. Explain how we have shown that $\det(A)$ is the product of the eigenvalues of A .

- (3) Find the eigenvalues of the following matrices. For each eigenvalue, determine its algebraic and geometric multiplicity.

(a) $A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$

(b) $A = \begin{bmatrix} 2 & 0 & 3 \\ 0 & 1 & 0 \\ 0 & 1 & 2 \end{bmatrix}$

- (4) Let A be an $n \times n$ matrix. Use the characteristic equation to explain why A and A^T have the same eigenvalues.

²This result is known as the Cayley-Hamilton Theorem and is one of the fascinating results in linear algebra.



- (5) Find three 3×3 matrices whose eigenvalues are 2 and 3, and for which the dimensions of the eigenspaces for $\lambda = 2$ and $\lambda = 3$ are different.
- (6) Suppose A is an $n \times n$ matrix and B is an invertible $n \times n$ matrix. Explain why the characteristic polynomial of A is the same as the characteristic polynomial of BAB^{-1} , and hence, as a result, the eigenvalues of A and BAB^{-1} are the same.
- (7) Label each of the following statements as True or False. Provide justification for your response.
 - (a) **True/False** If the determinant of a 2×2 matrix A is positive, then A has two distinct real eigenvalues.
 - (b) **True/False** If two 2×2 matrices have the same eigenvalues, then they have the same eigenvectors.
 - (c) **True/False** The characteristic polynomial of an $n \times n$ matrix has degree n .
 - (d) **True/False** If R is the reduced row echelon form of an $n \times n$ matrix A , then A and R have the same eigenvalues.
 - (e) **True/False** If R is the reduced row echelon form of an $n \times n$ matrix A , and \mathbf{v} is an eigenvector of A , then \mathbf{v} is an eigenvector of R .
 - (f) **True/False** Let A and B be $n \times n$ matrices with characteristic polynomials $p_A(\lambda)$ and $p_B(\lambda)$, respectively. If $A \neq B$, then $p_A(\lambda) \neq p_B(\lambda)$.
 - (g) **True/False** Every matrix has at least one eigenvalue.
 - (h) **True/False** Suppose A is a 3×3 matrix with three distinct eigenvalues. Then any three eigenvectors, one for each eigenvalue, will form a basis of \mathbb{R}^3 .
 - (i) **True/False** If an eigenvalue λ is repeated 3 times among the eigenvalues of a matrix, then there are at most 3 linearly independent eigenvectors corresponding to λ .

Project: The Ehrenfest Model

To realistically model the diffusion of gas molecules we would need to consider a system with a large number of balls as substitutes for the gas molecules. However, the main idea can be seen in a model with a much smaller number of balls, as we will do now. Suppose we have two bins that contain a total of 4 balls between them. Label the bins as Bin 1 and Bin 2. In this case we can think of entropy as the number of different possible ways the balls can be arranged in the system. For example, there is only 1 way for all of the balls to be in Bin 1 (low entropy), but there are 4 ways that we can have one ball in Bin 1 (choose any one of the four different balls, which can be distinguished from each other) and 3 balls in Bin 2 (higher entropy). The highest entropy state has the balls equally distributed between the bins (with 6 different ways to do this).

We assume that there is a way for balls to move from one bin to the other (like having gas molecules pass through a permeable membrane). A way to think about this is that we select a ball (from ball 1 to ball 4, which are different balls) and move that ball from its current bin to the other bin. Consider a “move” to be any instance when a ball changes bins. A *state* is any configuration of



balls in the bins at a given time, and the state changes when a ball is chosen at random and moved to the other bin. The possible states are to have 0 balls in Bin 1 and 4 balls in Bin 2 (State 0, entropy 1), 1 ball in Bin 1 and 3 in Bin 2 (State 1, entropy 4), 2 balls in each Bin (State 2, entropy 6), 3 balls in Bin 1 and 1 ball in Bin 2 (State 3, entropy 4), and 4 balls in Bin 1 and 0 balls in Bin 2 (State 4, entropy 1). These states are shown in Figure 17.3.

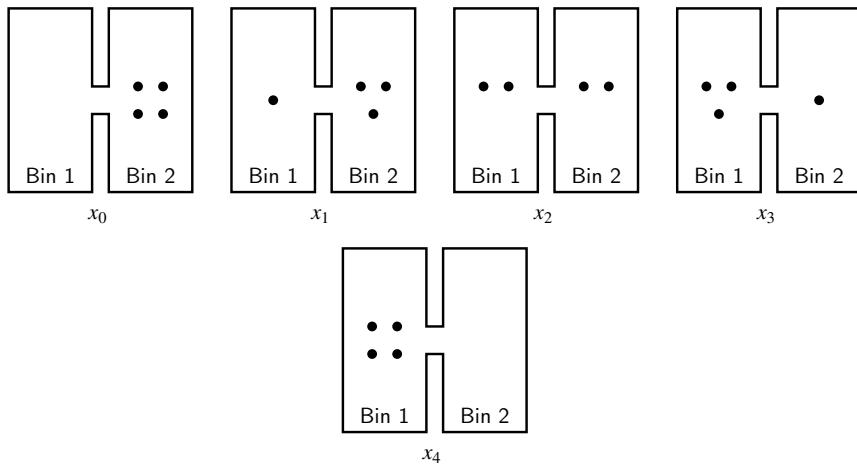


Figure 17.3: States

Project Activity 17.1. To model the system of balls in bins we need to understand how the system can transform from one state to another. It suffices to count the number of balls in Bin 1 (since the remaining balls will be in Bin 2). Even though the balls are labeled, our count only cares about how many balls are in each bin. Let $\mathbf{x}_0 = [x_0, x_1, x_2, x_3, x_4]^\top$, where x_i is the probability that Bin 1 contains i balls, and let $\mathbf{x}_1 = [x_0^1, x_1^1, x_2^1, x_3^1, x_4^1]^\top$, where x_i^1 is the probability that Bin 1 contains i balls after the first move. We will call the vectors \mathbf{x}_0 and \mathbf{x}_1 *probability distributions* of balls in bins. Note that since all four balls have to be placed in some bin, the sum of the entries in our probability distribution vectors must be 1. Recall that a move is an instance when a ball changes bins. We want to understand how \mathbf{x}_1 is obtained from \mathbf{x}_0 . In other words, we want to figure out what the probability that Bin 1 contains 0, 1, 2, 3, or 4 balls after one ball changes bins if our initial probability distribution of balls in bins is \mathbf{x}_0 .

We begin by analyzing the ways that a state can change. For example,

- Suppose there are 0 balls in Bin 1. (In our probability distribution \mathbf{x}_0 , this happens with probability x_0 .) Then there are four balls in Bin 2. The only way for a ball to change bins is if one of the four balls moves from Bin 2 to Bin 1, putting us in State 1. Regardless of which ball moves, we will always be put in State 1, so this happens with a probability of 1. In other words, if the probability that Bin 1 contains 0 balls is x_0 , then there is a probability of $(1)x_0$ that Bin 1 will contain 1 ball after the move.
- Suppose we have 1 ball in Bin 1. There are four ways this can happen (since there are four balls, and the one in Bin 1 is selected at random from the four balls), so the probability of a given ball being in Bin 1 is $\frac{1}{4}$.

- If the ball in Bin 1 moves, that move puts us in State 0. In other words, if the probability that Bin 1 contains 1 ball is x_1 , then there is a probability of $\frac{1}{4}x_1$ that Bin 1 will contain 0 balls after a move.
 - If any of the 3 balls in Bin 2 moves (each moves with probability $\frac{3}{4}$), that move puts us in State 2. In other words, if the probability that Bin 1 contains 1 ball is x_1 , then there is a probability of $\frac{3}{4}x_1$ that Bin 1 will contain 2 balls after a move.
- (a) Complete this analysis to explain the probabilities if there are 2, 3, or 4 balls in Bin 1.
- (b) Explain how the results of part (a) show that

$$\begin{aligned}x_0^1 &= 0x_0 + \frac{1}{4}x_1 + 0x_2 + 0x_3 + 0x_4 \\x_1^1 &= 1x_0 + 0x_1 + \frac{1}{2}x_2 + 0x_3 + 0x_4 \\x_2^1 &= 0x_0 + \frac{3}{4}x_1 + 0x_2 + \frac{3}{4}x_3 + 0x_4 \\x_3^1 &= 0x_0 + 0x_1 + \frac{1}{2}x_2 + 0x_3 + 1x_4 \\x_4^1 &= 0x_0 + 0x_1 + 0x_2 + \frac{1}{4}x_3 + 0x_4\end{aligned}$$

The system we developed in Project Activity 17.1 has matrix form

$$\mathbf{x}_1 = T\mathbf{x}_0,$$

where T is the *transition matrix*

$$T = \begin{bmatrix} 0 & \frac{1}{4} & 0 & 0 & 0 \\ 1 & 0 & \frac{1}{2} & 0 & 0 \\ 0 & \frac{3}{4} & 0 & \frac{3}{4} & 0 \\ 0 & 0 & \frac{1}{2} & 0 & 1 \\ 0 & 0 & 0 & \frac{1}{4} & 0 \end{bmatrix}.$$

Subsequent moves give probability distribution vectors

$$\begin{aligned}\mathbf{x}_2 &= T\mathbf{x}_1 \\ \mathbf{x}_3 &= T\mathbf{x}_2 \\ &\vdots && \vdots \\ \mathbf{x}_k &= T\mathbf{x}_{k-1}.\end{aligned}$$

This example is an example of a Markov process (see Definition 9.4). There are several questions we can ask about this model. For example, what is the long-term behavior of this system, and how does this model relate to entropy? That is, given an initial probability distribution vector \mathbf{x}_0 , the system will have probability distribution vectors $\mathbf{x}_1, \mathbf{x}_2, \dots$ after subsequent moves. What happens to the vectors \mathbf{x}_k as k goes to infinity, and what does this tell us about entropy? To answer these questions, we will first explore the sequence $\{\mathbf{x}_k\}$ numerically, and then use the eigenvalues and eigenvectors of T to analyze the sequence $\{\mathbf{x}_k\}$.

Project Activity 17.2. Use appropriate technology to do the following.



- (a) Suppose we begin with a probability distribution vector $\mathbf{x}_0 = [1 \ 0 \ 0 \ 0 \ 0]^T$. Calculate vectors \mathbf{x}_k for enough values of k so that you can identify the long term behavior of the sequence. Describe this behavior.

- (b) Repeat part (a) with

$$\text{i. } \mathbf{x}_0 = \left[0 \ \frac{1}{2} \ \frac{1}{2} \ 0 \ 0 \right]^T$$

$$\text{ii. } \mathbf{x}_0 = \left[0 \ \frac{1}{3} \ \frac{1}{3} \ 0 \ \frac{1}{3} \right]^T$$

$$\text{iii. } \mathbf{x}_0 = \left[\frac{1}{5} \ \frac{1}{5} \ \frac{1}{5} \ \frac{1}{5} \ \frac{1}{5} \right]^T$$

Describe the long term behavior of the sequence $\{\mathbf{x}_k\}$ in each case.

In what follows, we investigate the behavior of the sequence $\{\mathbf{x}_k\}$ that we uncovered in Project Activity 17.2.

Project Activity 17.3. We use the characteristic polynomial to find the eigenvalues of T .

- (a) Find the characteristic polynomial of T . Factor the characteristic polynomial into a product of linear polynomials to show that the eigenvalues of T are $0, 1, -1, \frac{1}{2}$ and $-\frac{1}{2}$.
- (b) As we will see a bit later, certain eigenvectors for T will describe the end behavior of the sequence $\{\mathbf{x}_k\}$. Find eigenvectors for T corresponding to the eigenvalues 1 and -1 . Explain how the eigenvector for T corresponding to the eigenvalue 1 explains the behavior of one of the sequences we saw in Project Activity 17.2. (Any eigenvector of T with eigenvalue 1 is called an *equilibrium* or *steady state* vector.)

Now we can analyze the behavior of the sequence $\{\mathbf{x}_k\}$.

Project Activity 17.4. To make the notation easier, we will let \mathbf{v}_1 be an eigenvector of T corresponding to the eigenvalue 0 , \mathbf{v}_2 an eigenvector of T corresponding to the eigenvalue 1 , \mathbf{v}_3 an eigenvector of T corresponding to the eigenvalue -1 , \mathbf{v}_4 an eigenvector of T corresponding to the eigenvalue $\frac{1}{2}$, and \mathbf{v}_5 an eigenvector of T corresponding to the eigenvalue $-\frac{1}{2}$.

- (a) Explain why $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4, \mathbf{v}_5\}$ is a basis of \mathbb{R}^5 .
- (b) Let \mathbf{x}_0 be any initial probability distribution vector. Explain why we can write \mathbf{x}_0 as

$$\mathbf{x}_0 = a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + a_3\mathbf{v}_3 + a_4\mathbf{v}_4 + a_5\mathbf{v}_5 = \sum_{i=1}^5 a_i\mathbf{v}_i$$

for some scalars a_1, a_2, a_3, a_4 , and a_5 .

We can now use the eigenvalues and eigenvectors of T to write the vectors \mathbf{x}_k in a convenient form. Let $\lambda_1 = 0$, $\lambda_2 = 1$, $\lambda_3 = -1$, $\lambda_4 = \frac{1}{2}$, and $\lambda_5 = -\frac{1}{2}$. Notice that

$$\begin{aligned} \mathbf{x}_1 &= T\mathbf{x}_0 \\ &= T(a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + a_3\mathbf{v}_3 + a_4\mathbf{v}_4 + a_5\mathbf{v}_5) \\ &= a_1T\mathbf{v}_1 + a_2T\mathbf{v}_2 + a_3T\mathbf{v}_3 + a_4T\mathbf{v}_4 + a_5T\mathbf{v}_5 \\ &= a_1\lambda_1\mathbf{v}_1 + a_2\lambda_2\mathbf{v}_2 + a_3\lambda_3\mathbf{v}_3 + a_4\lambda_4\mathbf{v}_4 + a_5\lambda_5\mathbf{v}_5 \\ &= \sum_{i=1}^5 a_i\lambda_i\mathbf{v}_i. \end{aligned}$$



Similarly

$$\mathbf{x}_2 = T\mathbf{x}_1 = T \left(\sum_{i=1}^5 a_i \lambda_i \mathbf{v}_i \right) = \sum_{i=1}^5 a_i \lambda_i T \mathbf{v}_i = \sum_{i=1}^5 a_i \lambda_i^2 \mathbf{v}_i.$$

We can continue in this manner to ultimately show that for each positive integer k we have

$$\mathbf{x}_k = \sum_{i=1}^5 a_i \lambda_i^k \mathbf{v}_i \quad (17.4)$$

when $\mathbf{x}_0 = \sum_{i=1}^5 a_i \mathbf{v}_i$.

Project Activity 17.5. Recall that we are interested in understanding the behavior of the sequence $\{\mathbf{x}_k\}$ as k goes to infinity.

- (a) Equation (17.4) shows that we need to know $\lim_{k \rightarrow \infty} \lambda_i^k$ for each i in order to analyze $\lim_{k \rightarrow \infty} \mathbf{x}_k$. Calculate or describe these limits.
- (b) Use the result of part (a), Equation (17.4), and Project Activity 17.3 (b) to explain why the sequence $\{\mathbf{x}_k\}$ is either eventually fixed or oscillates between two states. Compare to the results from Project Activity 17.2. How are these results related to entropy? You may use the facts that
 - $\mathbf{v}_1 = [1 \ 0 \ -2 \ 0 \ 1]^T$ is an eigenvector for T corresponding to the eigenvalue 0,
 - $\mathbf{v}_2 = [1 \ 4 \ 6 \ 4 \ 1]^T$ is an eigenvector for T corresponding to the eigenvalue 1,
 - $\mathbf{v}_3 = [1 \ -4 \ 6 \ -4 \ 1]^T$ is an eigenvector for T corresponding to the eigenvalue -1 ,
 - $\mathbf{v}_4 = [-1 \ -2 \ 0 \ 2 \ 1]^T$ is an eigenvector for T corresponding to the eigenvalue $\frac{1}{2}$,
 - $\mathbf{v}_5 = [-1 \ 2 \ 0 \ -2 \ 1]^T$ is an eigenvector for T corresponding to the eigenvalue $-\frac{1}{2}$.

Section 18

Diagonalization

Focus Questions

By the end of this section, you should be able to give precise and thorough answers to the questions listed below. You may want to keep these questions in mind to focus your thoughts as you complete the section.

- What is a diagonal matrix?
- What does it mean to diagonalize a matrix?
- What does it mean for two matrices to be similar?
- What important properties do similar matrices share?
- Under what conditions is a matrix diagonalizable?
- When a matrix A is diagonalizable, what is the structure of a matrix P that diagonalizes A ?
- Why is diagonalization useful?

Application: The Fibonacci Numbers

In 1202 Leonardo of Pisa (better known as Fibonacci) published *Liber Abaci* (roughly translated as *The Book of Calculation*), in which he constructed a mathematical model of the growth of a rabbit population. The problem Fibonacci considered is that of determining the number of pairs of rabbits produced in a given time period beginning with an initial pair of rabbits. Fibonacci made the assumptions that each pair of rabbits more than one month old produces a new pair of rabbits each month, and that no rabbits die. (We ignore any issues about that might arise concerning the gender of the offspring.) If we let F_n represent the number of rabbits in month n , Fibonacci produced the model

$$F_{n+2} = F_{n+1} + F_n, \quad (18.1)$$

for $n \geq 0$ where $F_0 = 0$ and $F_1 = 1$. The resulting sequence

$$1, 1, 2, 3, 5, 8, 13, 21, \dots$$

is a very famous sequence in mathematics and is called the Fibonacci sequence. This sequence is thought to model many natural phenomena such as number of seeds in a sunflower and anything which grows in a spiral form. It is so famous in fact that it has a journal devoted entirely to it. As a note, while Fibonacci's work *Liber Abaci* introduced this sequence to the western world, it had been described earlier Sanskrit texts going back as early as the sixth century.

By definition, the Fibonacci numbers are calculated by recursion. This is a very ineffective way to determine entries F_n for large n . Later in this section we will derive a fascinating and unexpected formula for the Fibonacci numbers using the idea of diagonalization.

Introduction

As we have seen when studying Markov processes, each state is dependent on the previous state. If \mathbf{x}_0 is the initial state and A is the transition matrix, then the n th state is found by $A^n \mathbf{x}_0$. In these situations, and others, it is valuable to be able to quickly and easily calculate powers of a matrix. We explore a way to do that in this section.

Preview Activity 18.1. Consider a very simplified weather forecast. Let us assume there are two possible states for the weather: rainy (R) or sunny(S). Let us also assume that the weather patterns are stable enough that we can reasonably predict the weather tomorrow based on the weather today. If it is sunny today, then there is a 70% chance that it will be sunny tomorrow, and if it is rainy today then there is a 40% chance that it will be rainy tomorrow. If $\mathbf{x}_0 = \begin{bmatrix} s \\ r \end{bmatrix}$ is a state vector that indicates a probability s that it is sunny and probability r that it is rainy on day 0, then

$$\mathbf{x}_1 = \begin{bmatrix} 0.70 & 0.40 \\ 0.30 & 0.60 \end{bmatrix} \mathbf{x}_0$$

tells us the likelihood of it being sunny or rainy on day 1. Let $A = \begin{bmatrix} 0.70 & 0.40 \\ 0.30 & 0.60 \end{bmatrix}$.

- (1) Suppose it is sunny today, that is $\mathbf{x}_0 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$. Calculate $\mathbf{x}_1 = A\mathbf{x}_0$ and explain how this matrix-vector product tells us the probability that it will be sunny tomorrow.
- (2) Calculate $\mathbf{x}_2 = A\mathbf{x}_1$ and interpret the meaning of each component of the product.
- (3) Explain why $\mathbf{x}_2 = A^2\mathbf{x}_0$. Then explain in general why $\mathbf{x}_n = A^n\mathbf{x}_0$.
- (4) The previous result demonstrates that to determine the long-term probability of a sunny or rainy day, we want to be able to easily calculate powers of the matrix A . Use a computer algebra system (e.g., Maple, Mathematica, Wolfram|Alpha) to calculate the entries of \mathbf{x}_{10} , \mathbf{x}_{20} , and \mathbf{x}_{30} . Based on this data, what do you expect the long term probability of any day being a sunny one?



Diagonalization

In Preview Activity 18.1 we saw how if we can powers of a matrix we can make predictions about the long-term behavior of some systems. In general, calculating powers of a matrix can be a very difficult thing, but there are times when the process is straightforward.

Activity 18.1. Let $D = \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix}$.

- (a) Show that $D^2 = \begin{bmatrix} 2^2 & 0 \\ 0 & 3^2 \end{bmatrix}$.
- (b) Show that $D^3 = \begin{bmatrix} 2^3 & 0 \\ 0 & 3^3 \end{bmatrix}$. (Hint: $D^3 = DD^2$.)
- (c) Explain in general why $D^n = \begin{bmatrix} 2^n & 0 \\ 0 & 3^n \end{bmatrix}$ for any positive integer n .

Activity 18.1 illustrates that calculating powers of square matrices whose only nonzero entries are along the diagonal is rather simple. In general, if

$$D = \begin{bmatrix} d_{11} & 0 & 0 & \cdots & 0 & 0 \\ 0 & d_{22} & 0 & \cdots & 0 & 0 \\ \vdots & 0 & 0 & \ddots & & \vdots \\ 0 & 0 & 0 & \cdots & 0 & d_{nn} \end{bmatrix},$$

then

$$D^k = \begin{bmatrix} d_{11}^k & 0 & 0 & \cdots & 0 & 0 \\ 0 & d_{22}^k & 0 & \cdots & 0 & 0 \\ \vdots & 0 & 0 & \ddots & & \vdots \\ 0 & 0 & 0 & \cdots & 0 & d_{nn}^k \end{bmatrix}$$

for any positive integer k . Recall that a diagonal matrix is a matrix whose only nonzero elements are along the diagonal (see Definition 8.6). In this section we will see that matrices that are similar to diagonal matrices have some very nice properties, and that diagonal matrices are useful in calculations of powers of matrices.

We can utilize the method of calculating powers of diagonal matrices to also easily calculate powers of other types of matrices.

Activity 18.2. Let D be any matrix, P an invertible matrix, and let $A = P^{-1}DP$.

- (a) Show that $A^2 = P^{-1}D^2P$.
- (b) Show that $A^3 = P^{-1}D^3P$.
- (c) Explain in general why $A^n = P^{-1}D^nP$ for positive integers n .

As Activity 18.2 illustrates, to calculate the powers of a matrix of the form $P^{-1}DP$ we only need determine the powers of the matrix D . If D is a diagonal matrix, this is especially straightforward.

Similar Matrices

Similar matrices play an important role in certain calculations. For example, Activity 18.2 showed that if we can write a square matrix A in the form $A = P^{-1}DP$ for some invertible matrix P and diagonal matrix D , then finding the powers of A is straightforward. As we will see, the relation $A = P^{-1}DP$ will imply that the matrices A and D share many properties.

Definition 18.1. The $n \times n$ matrix A is **similar** to the $n \times n$ matrix B if there is an invertible matrix P such that $A = P^{-1}BP$.

Activity 18.3. Let $A = \begin{bmatrix} 1 & 1 \\ 2 & 0 \end{bmatrix}$ and $B = \begin{bmatrix} 2 & 2 \\ 0 & -1 \end{bmatrix}$. Assume that A is similar to B via the matrix $P = \begin{bmatrix} 2 & 1 \\ 2 & 2 \end{bmatrix}$.

- Calculate $\det(A)$ and $\det(B)$. What do you notice?
- Find the characteristic polynomials of A and B . What do you notice?
- What can you say about the eigenvalues of A and B ? Explain.
- Explain why $\mathbf{x} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ is an eigenvector for A with eigenvalue 2. Is \mathbf{x} an eigenvector for B with eigenvalue 2? Why or why not?

Activity 18.3 suggests that similar matrices share some, but not all, properties. Note that if $A = P^{-1}BP$, then $B = Q^{-1}AQ$ with $Q = P^{-1}$. So if A is similar to B , then B is similar to A . Similarly (no pun intended), since $A = I^{-1}AI$ (where I is the identity matrix), then any square matrix is similar to itself. Also, if $A = P^{-1}BP$ and $B = M^{-1}CM$, then $A = (MP)^{-1}C(MP)$. So if A is similar to B and B is similar to C , then A is similar to C . If you have studied relations, these three properties show that similarity is an equivalence relation on the set of all $n \times n$ matrices. This is one reason why similar matrices share many important traits, as the next activity highlights.

Activity 18.4. Let A and B be similar matrices with $A = P^{-1}BP$.

- Use the multiplicative property of the determinant to explain why $\det(A) = \det(B)$. So similar matrices have the same determinants.
- Use the fact that $P^{-1}IP = I$ to show that $A - \lambda I$ is similar to $B - \lambda I$.
- Explain why it follows from (a) and (b) that

$$\det(A - \lambda I) = \det(B - \lambda I).$$

So similar matrices have the same characteristic polynomial, and the same eigenvalues.

We summarize some properties of similar matrices in the following theorem.

Theorem 18.2. Let A and B be similar $n \times n$ matrices and I the $n \times n$ identity matrix. Then

- $\det(A) = \det(B)$,



- (2) $A - \lambda I$ is similar to $B - \lambda I$,
- (3) A and B have the same characteristic polynomial,
- (4) A and B have the same eigenvalues.

Similarity and Matrix Transformations

When a matrix is similar to a diagonal matrix, we can gain insight into the action of the corresponding matrix transformation. As an example, consider the matrix transformation T from \mathbb{R}^2 to \mathbb{R}^2 defined by $T\mathbf{x} = A\mathbf{x}$, where

$$A = \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix}. \quad (18.2)$$

We are interested in understanding what this matrix transformation does to vectors in \mathbb{R}^2 . First we note that A has eigenvalues $\lambda_1 = 2$ and $\lambda_2 = 4$ with corresponding eigenvectors $\mathbf{v}_1 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$ and $\mathbf{v}_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$. If we let $P = [\mathbf{v}_1 \ \mathbf{v}_2]$, then you can check that

$$P^{-1}AP = D$$

and

$$A = PDP^{-1},$$

where

$$D = \begin{bmatrix} 2 & 0 \\ 0 & 4 \end{bmatrix}.$$

Thus,

$$T(\mathbf{x}) = PDP^{-1}\mathbf{x}.$$

A simple calculation shows that

$$P^{-1} = \frac{1}{2} \begin{bmatrix} -1 & 1 \\ 1 & 1 \end{bmatrix}.$$

Let us apply T to the unit square whose sides are formed by the vectors $\mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $\mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ as shown in the first picture in Figure 18.1.

To apply T we first multiply \mathbf{e}_1 and \mathbf{e}_2 by P^{-1} . This gives us

$$P^{-1}\mathbf{e}_1 = \frac{1}{2}\mathbf{v}_1 \text{ and } P^{-1}\mathbf{e}_2 = \frac{1}{2}\mathbf{v}_2.$$

So P^{-1} transforms the standard coordinate system into a coordinate system in which \mathbf{v}_1 and \mathbf{v}_2 determine the axes, as illustrated in the second picture in Figure 18.1. Applying D to the output scales by 2 in the \mathbf{v}_1 direction and 4 in the \mathbf{v}_2 direction as depicted in the third picture in Figure 18.1. Finally, we apply P to translate back into the standard xy coordinate system as shown in the last picture in Figure 18.1.

This example illustrates that it is most convenient to view the action of T in the coordinate system where \mathbf{v}_1 serves as the x -direction and \mathbf{v}_2 as the y -direction. In this case, we can visualize that when we apply the transformation T to a vector in this system it is just scaled in both directions by the matrix D . Then the matrix P translates everything back to the standard xy coordinate system.

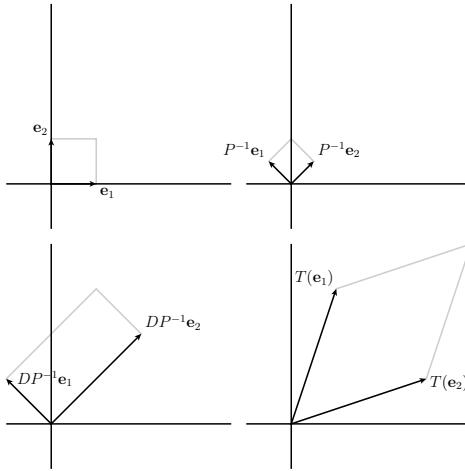


Figure 18.1: The matrix transformation.

This geometric perspective provides another example of how having a matrix similar to a diagonal matrix informs us about the situation. In what follows we determine the conditions that determine when a matrix is similar to a diagonal matrix.

Diagonalization in General

In Preview Activity 18.1 and in the matrix transformation example we found that a matrix A was similar to a diagonal matrix whose columns were eigenvectors of A . This will work for a general $n \times n$ matrix A as long as we can find an invertible matrix P whose columns are eigenvectors of A . More specifically, suppose A is an $n \times n$ matrix with n linearly independent eigenvectors $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$ with corresponding eigenvalues $\lambda_1, \lambda_2, \dots, \lambda_n$ (not necessarily distinct). Let

$$P = [\mathbf{v}_1 \ \mathbf{v}_2 \ \mathbf{v}_3 \ \cdots \ \mathbf{v}_n].$$

Then

$$\begin{aligned}
 AP &= [A\mathbf{v}_1 \ A\mathbf{v}_2 \ A\mathbf{v}_3 \ \cdots \ A\mathbf{v}_n] \\
 &= [\lambda_1\mathbf{v}_1 \ \lambda_2\mathbf{v}_2 \ \lambda_3\mathbf{v}_3 \ \cdots \ \lambda_n\mathbf{v}_n] \\
 &= [\mathbf{v}_1 \ \mathbf{v}_2 \ \mathbf{v}_3 \ \cdots \ \mathbf{v}_n] \begin{bmatrix} \lambda_1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & \lambda_2 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & & \cdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda_{n-1} & 0 \\ 0 & 0 & 0 & \cdots & 0 & \lambda_n \end{bmatrix} \\
 &= PD.
 \end{aligned}$$

where

$$D = \begin{bmatrix} \lambda_1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & \lambda_2 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & & \cdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda_{n-1} & 0 \\ 0 & 0 & 0 & \cdots & 0 & \lambda_n \end{bmatrix}.$$

Since the columns of P are linearly independent, we know P is invertible, and so

$$P^{-1}AP = D.$$

Definition 18.3. An $n \times n$ matrix A is **diagonalizable** if there is an invertible $n \times n$ matrix P so that $P^{-1}AP$ is a diagonal matrix.

In other words, a matrix A is diagonalizable if A is similar to a diagonal matrix.

IMPORTANT NOTE: The key notion to the process described above is that in order to diagonalize an $n \times n$ matrix A , we have to find n linearly independent eigenvectors for A . When A is diagonalizable, a matrix P so that $P^{-1}AP$ is diagonal is said to *diagonalize* A .

Activity 18.5. Find an invertible matrix P that diagonalizes A .

$$(a) \ A = \begin{bmatrix} 1 & 1 \\ 0 & 2 \end{bmatrix}$$

$$(b) \ A = \begin{bmatrix} 3 & 2 & 4 \\ 2 & 0 & 2 \\ 4 & 2 & 3 \end{bmatrix}. \text{ (Hint: The eigenvalues of } A \text{ are 8 and -1.)}$$

It should be noted that there are square matrices that are not diagonalizable. For example, the matrix $A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$ has 1 as its only eigenvalue and the dimension of the eigenspace of A corresponding to the eigenvalue is one. Therefore, it will be impossible to find two linearly independent eigenvectors for A .

We showed previously that eigenvectors corresponding to distinct eigenvalue are always linearly independent, so if an $n \times n$ matrix A has n distinct eigenvalues then A is diagonalizable. Activity 18.5 (b) shows that it is possible to diagonalize an $n \times n$ matrix even if the matrix does not have



n distinct eigenvalues. In general, we can diagonalize a matrix as long as the dimension of each eigenspace is equal to the multiplicity of the corresponding eigenvalue. In other words, a matrix is diagonalizable if the geometric multiplicity is the same as the algebraic multiplicity for each eigenvalue.

At this point we might ask one final question. We argued that if an $n \times n$ matrix A has n linearly independent eigenvectors, then A is diagonalizable. It is reasonable to wonder if the converse is true – that is, if A is diagonalizable, must A have n linearly independent eigenvectors? The answer is yes, and you are asked to show this in Exercise 6. We summarize the result in the following theorem.

Theorem 18.4 (The Diagonalization Theorem). *An $n \times n$ matrix A is diagonalizable if and only if A has n linearly independent eigenvectors. If A is diagonalizable and has linearly independent eigenvectors $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$ with $A\mathbf{v}_i = \lambda_i \mathbf{v}_i$ for each i , then $n \times n$ matrix $P[\mathbf{v}_1 \ \mathbf{v}_2 \ \cdots \ \mathbf{v}_n]$ whose columns are linearly independent eigenvectors of A satisfies $P^{-1}AP = D$, where $D[d_{ij}]$ is the diagonal matrix with diagonal entries $d_{ii} = \lambda_i$ for each i .*

Examples

What follows are worked examples that use the concepts from this section.

Example 18.5. Let $A = \begin{bmatrix} 1 & -2 & 1 \\ 0 & 3 & -1 \\ 0 & -2 & 2 \end{bmatrix}$ and $B = \begin{bmatrix} 1 & 2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 4 \end{bmatrix}$. You should use appropriate technology to calculate determinants, perform any row reductions, or solve any polynomial equations.

- Determine if A is diagonalizable. If diagonalizable, find a matrix P that diagonalizes A .
- Determine if B is diagonalizable. If diagonalizable, find a matrix Q that diagonalizes B .
- Is it possible for two matrices R and S to have the same eigenvalues with the same algebraic multiplicities, but one matrix is diagonalizable and the other is not? Explain.

Example Solution.

- Technology shows that the characteristic polynomial of A is

$$p(\lambda) = \det(A - \lambda I_3) = (4 - \lambda)(1 - \lambda)^2.$$

The eigenvalues of A are the solutions to the characteristic equation $p(\lambda) = 0$. Thus, the eigenvalues of A are 1 and 4.

To find a basis for the eigenspace of A corresponding to the eigenvalue 1, we find the general solution to the homogeneous system $(A - I_3)\mathbf{x} = \mathbf{0}$. Using technology we see

that the reduced row echelon form of $A - I_3 = \begin{bmatrix} 0 & -2 & 1 \\ 0 & 2 & -1 \\ 0 & -2 & 1 \end{bmatrix}$ is $\begin{bmatrix} 0 & 1 & -\frac{1}{2} \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$. So



if $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$, then the general solution to $(A - I_3)\mathbf{x} = \mathbf{0}$ is

$$\begin{aligned}\mathbf{x} &= \begin{bmatrix} x_1 \\ \frac{1}{2}x_3 \\ x_3 \end{bmatrix} \\ &= x_1 \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} 0 \\ \frac{1}{2} \\ 1 \end{bmatrix}.\end{aligned}$$

So a basis for the eigenspace of A corresponding to the eigenvalue 1 is

$$\left\{ [1 \ 0 \ 0]^T, [0 \ 1 \ 2]^T \right\}.$$

To find a basis for the eigenspace of A corresponding to the eigenvalue 4, we find the general solution to the homogeneous system $(A - 4I_3)\mathbf{x} = \mathbf{0}$. Using technology we see

that the reduced row echelon form of $A - 4I_3 = \begin{bmatrix} -3 & -2 & 1 \\ 0 & -1 & -1 \\ 0 & -2 & -2 \end{bmatrix}$ is $\begin{bmatrix} 0 & 1 & -1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix}$.

So if $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$, then the general solution to $(A - 4I_3)\mathbf{x} = \mathbf{0}$ is

$$\begin{aligned}\mathbf{x} &= [x_1 \ x_2 \ x_3]^T \\ &= [x_3 \ -x_3 \ x_3]^T \\ &= x_3 [1 \ -1 \ 1]^T.\end{aligned}$$

So a basis for the eigenspace of A corresponding to the eigenvalue 4 is

$$\left\{ [1 \ -1 \ 0]^T \right\}.$$

Eigenvectors corresponding to different eigenvalues are linearly independent, so the set

$$\left\{ [1 \ 0 \ 0]^T, [0 \ 1 \ 2]^T, [1 \ -1 \ 0]^T \right\}$$

is a basis for \mathbb{R}^3 . Since we can find a basis for \mathbb{R}^3 consisting of eigenvectors of A , we conclude that A is diagonalizable. Letting

$$P = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & -1 \\ 0 & 2 & 1 \end{bmatrix}$$

gives us

$$P^{-1}AP = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 4 \end{bmatrix}$$



- (b) Technology shows that the characteristic polynomial of B is

$$p(\lambda) = \det(B - \lambda I_3) = (4 - \lambda)(1 - \lambda)^2.$$

The eigenvalues of B are the solutions to the characteristic equation $p(\lambda) = 0$. Thus, the eigenvalues of B are 1 and 4.

To find a basis for the eigenspace of B corresponding to the eigenvalue 1, we find the general solution to the homogeneous system $(B - I_3)\mathbf{x} = \mathbf{0}$. Using technology we see

that the reduced row echelon form of $B - I_3 = \begin{bmatrix} 0 & 2 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 3 \end{bmatrix}$ is $\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$. So if

$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$, then the general solution to $(B - I_3)\mathbf{x} = \mathbf{0}$ is

$$\begin{aligned} \mathbf{x} &= [x_1 \ x_2 \ x_3]^\top \\ &= [x_1 \ 0 \ 0]^\top \\ &= x_1 [1 \ 0 \ 0]^\top. \end{aligned}$$

So a basis for the eigenspace of B corresponding to the eigenvalue 1 is

$$\{[1 \ 0 \ 0]^\top\}.$$

To find a basis for the eigenspace of B corresponding to the eigenvalue 4, we find the general solution to the homogeneous system $(B - 4I_3)\mathbf{x} = \mathbf{0}$. Using technology we see

that the reduced row echelon form of $B - 4I_3 = \begin{bmatrix} -3 & 2 & 0 \\ 0 & -3 & 0 \\ 0 & 0 & 0 \end{bmatrix}$ is $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$. So if

$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$, then the general solution to $(B - 4I_3)\mathbf{x} = \mathbf{0}$ is

$$\begin{aligned} \mathbf{x} &= [x_1 \ x_2 \ x_3]^\top \\ &= [0 \ 0 \ x_3]^\top \\ &= x_3 [0 \ 0 \ 1]^\top. \end{aligned}$$

So a basis for the eigenspace of B corresponding to the eigenvalue 4 is

$$\{[0 \ 0 \ 1]^\top\}.$$

Since each eigenspace is one-dimensional, we cannot find a basis for \mathbb{R}^3 consisting of eigenvectors of B . We conclude that B is not diagonalizable.

- (c) Yes it is possible for two matrices R and S to have the same eigenvalues with the same multiplicities, but one matrix is diagonalizable and the other is not. An example is given by the matrices A and B in this problem.



Example 18.6.

- (a) Is it possible to find diagonalizable matrices A and B such that AB is not diagonalizable? If yes, provide an example. If no, explain why.
- (b) Is it possible to find diagonalizable matrices A and B such that $A+B$ is not diagonalizable? If yes, provide an example. If no, explain why.
- (c) Is it possible to find a diagonalizable matrix A such that A^T is not diagonalizable? If yes, provide an example. If no, explain why.
- (d) Is it possible to find an invertible diagonalizable matrix A such that A^{-1} is not diagonalizable? If yes, provide an example. If no, explain why.

Example Solution.

- (a) Let $A = \begin{bmatrix} 1 & 1 \\ 0 & 2 \end{bmatrix}$ and $B = \begin{bmatrix} 2 & -2 \\ 0 & 1 \end{bmatrix}$. Since A and B are both diagonal matrices, their eigenvalues are their diagonal entries. With 2 distinct eigenvalues, both A and B are diagonalizable. In this case we have $AB = \begin{bmatrix} 2 & -1 \\ 0 & 2 \end{bmatrix}$, whose only eigenvector is 2. The reduced row echelon form of $AB - 2I_2$ is $\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$. So a basis for the eigenspace of AB is $\{[1 0]^T\}$. Since there is no basis for \mathbb{R}^2 consisting of eigenvectors of AB , we conclude that AB is not diagonalizable.
- (b) Let $A = \begin{bmatrix} 1 & 3 \\ 0 & 2 \end{bmatrix}$ and $B = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}$. Since A and B are both diagonal matrices, their eigenvalues are their diagonal entries. With 2 distinct eigenvalues, both A and B are diagonalizable. In this case we have $A + B = \begin{bmatrix} 3 & 3 \\ 0 & 3 \end{bmatrix}$, whose only eigenvector is 3. The reduced row echelon form of $(A + B) - 3I_2$ is $\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$. So a basis for the eigenspace of $A + B$ is $\{[1 0]^T\}$. Since there is no basis for \mathbb{R}^2 consisting of eigenvectors of $A + B$, we conclude that $A + B$ is not diagonalizable.
- (c) It is not possible to find a diagonalizable matrix A such that A^T is not diagonalizable. To see why, suppose that matrix A is diagonalizable. That is, there exists a matrix P such that $P^{-1}AP = D$, where D is a diagonal matrix. Recall that $(P^{-1})^T = (P^T)^{-1}$. So

$$\begin{aligned} D &= D^T \\ &= (P^{-1}AP)^T \\ &= P^T A^T (P^{-1})^T \\ &= P^T A^T (P^T)^{-1}. \end{aligned}$$

Letting $A = (P^T)^{-1}$, we conclude that

$$Q^{-1}A^TQ = D.$$



Therefore, Q diagonalizes A^T .

- (d) It is not possible to find an invertible diagonalizable matrix A such that A^{-1} is not diagonalizable. To see why, suppose that matrix A is diagonalizable. That is, there exists a matrix P such that $P^{-1}AP = D$, where D is a diagonal matrix. Thus, $A = PDP^{-1}$. Since A is invertible, $\det(A) \neq 0$. It follows that $\det(D) \neq 0$. So none of the diagonal entries of D can be 0. Thus, D is invertible and D^{-1} is a diagonal matrix. Then

$$D^{-1} = (P^{-1}AP)^{-1} = PA^{-1}P^{-1}$$

and so P^{-1} diagonalizes A^{-1} .

Summary

- A matrix $D = [d_{ij}]$ is a diagonal matrix if $d_{ij} = 0$ whenever $i \neq j$.
- A matrix A is diagonalizable if there is an invertible matrix P so that $P^{-1}AP$ is a diagonal matrix.
- Two matrices A and B are similar if there is an invertible matrix P so that

$$B = P^{-1}AP.$$

- Similar matrices have the same determinants, same characteristic polynomials, and same eigenvalues. Note that similar matrices do not necessarily have the same eigenvectors corresponding to the same eigenvalues.
- An $n \times n$ matrix A is diagonalizable if and only if A has n linearly independent eigenvectors.
- When an $n \times n$ matrix A is diagonalizable, then $P = [\mathbf{v}_1 \ \mathbf{v}_2 \ \mathbf{v}_3 \ \cdots \ \mathbf{v}_n]$ is invertible and $P^{-1}AP$ is diagonal, where $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$ are n linearly independent eigenvectors for A .
- One use for diagonalization is that once we have diagonalized a matrix A we can quickly and easily compute powers of A . Diagonalization can also help us understand the actions of matrix transformations.

Exercises

- (1) Determine if each of the following matrices is diagonalizable or not. For diagonalizable matrices, clearly identify a matrix P which diagonalizes the matrix, and what the resulting diagonal matrix is.

(a) $A = \begin{bmatrix} 2 & -1 \\ 1 & 4 \end{bmatrix}$

(b) $A = \begin{bmatrix} -1 & 4 & -2 \\ -3 & 4 & 0 \\ -3 & 1 & 3 \end{bmatrix}$



(2) The 3×3 matrix A has two eigenvalues $\lambda_1 = 2$ and $\lambda_2 = 3$. The vectors $\begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}$, $\begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix}$,

and $\begin{bmatrix} 2 \\ 4 \\ 2 \end{bmatrix}$ are eigenvectors for $\lambda_1 = 2$, while the vectors $\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$, $\begin{bmatrix} 2 \\ 2 \\ 2 \end{bmatrix}$ are eigenvectors for $\lambda_2 = 3$. Find the matrix A .

(3) Find a 2×2 non-diagonal matrix A and two different pairs of P and D matrices for which $A = PDP^{-1}$.

(4) Find a 2×2 non-diagonal matrix A and two different P matrices for which $A = PDP^{-1}$ with the same D .

(5) Suppose a 4×4 matrix A has eigenvalues 2, 3 and 5 and the eigenspace for the eigenvalue 3 has dimension 2. Do we have enough information to determine if A is diagonalizable? Explain.

(6) Let A be a diagonalizable $n \times n$ matrix. Show that A has n linearly independent eigenvectors.

(7)

(a) Let $A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$ and $B = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}$. Find the eigenvalues and eigenvectors of A and B . Conclude that it is possible for two different $n \times n$ matrices A and B to have exactly the same eigenvectors and corresponding eigenvalues.

(b) A natural question to ask is if there are any conditions under which $n \times n$ matrices that have exactly the same eigenvectors and corresponding eigenvalues must be equal. Determine the answer to this question if A and B are both diagonalizable.

(8)

(a) Show that if D and D' are $n \times n$ diagonal matrices, then $DD' = D'D$.

(b) Show that if A and B are $n \times n$ matrices and P is an invertible $n \times n$ matrix such that $P^{-1}AP = D$ and $P^{-1}BP = D'$ with D and D' diagonal matrices, then $AB = BA$.

(9) Exercise 2 in Section 17 shows that the determinant of a matrix is the product of its eigenvalues. In this exercise we show that the trace of a diagonalizable matrix is the sum of its eigenvalues.¹ First we define the trace of a matrix.

Definition 18.7. The **trace** of an $n \times n$ matrix $A = [a_{ij}]$ is the sum of the diagonal entries of A . That is,

$$\text{trace}(A) = a_{11} + a_{22} + \cdots + a_{nn} = \sum_{i=1}^n a_{ii}.$$

(a) Show that if $R = [r_{ij}]$ and $S = [s_{ij}]$ are $n \times n$ matrices, then $\text{trace}(RS) = \text{trace}(SR)$.

¹This result is true for any matrix, but the argument is more complicated.

- (b) Let A be a diagonalizable $n \times n$ matrix, and let $p(\lambda) = \det(A - \lambda I_n)$ be the characteristic polynomial of A . Let P be an invertible matrix such that $P^{-1}AP = D$, where D is the diagonal matrix whose diagonal entries are $\lambda_1, \lambda_2, \dots, \lambda_n$, the eigenvalues of A (note that these eigenvalues may not all be distinct).
- Explain why $\text{trace}(A) = \text{trace}(D)$.
 - Show that the trace of an $n \times n$ diagonalizable matrix is the sum of the eigenvalues of the matrix.
- (10) In this exercise we generalize the result of Exercise 12 in Section 8 to arbitrary diagonalizable matrices.

- (a) Show that if

$$D = \begin{bmatrix} \lambda_1 & 0 & 0 & \cdots & 0 \\ 0 & \lambda_2 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda_n \end{bmatrix},$$

then

$$e^D = \begin{bmatrix} e^{\lambda_1} & 0 & 0 & \cdots & 0 \\ 0 & e^{\lambda_2} & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & e^{\lambda_n} \end{bmatrix}.$$

- (b) Now suppose that an $n \times n$ matrix A is diagonalizable, with $P^{-1}AP$ equal to a diagonal matrix D . Show that $e^A = Pe^D P^{-1}$.
- (11) Let $A = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}$ and let $B = \begin{bmatrix} 0 & -1 \\ 0 & 0 \end{bmatrix}$.
- Use the result of Exercise 10 to calculate e^A .
 - Calculate e^B . (Hint: Explain why B is not diagonalizable.)
 - Use the result of Exercise 10 to calculate e^{A+B} .
 - The real exponential function satisfies some familiar properties. For example, $e^x e^y = e^y e^x$ and $e^{x+y} = e^x e^y$ for any real numbers x and y . Does the matrix exponential satisfy the corresponding properties. That is, if X and Y are $n \times n$ matrices, must $e^X e^Y = e^Y e^X$ and $e^{X+Y} = e^X e^Y$? Explain.
- (12) In Exercise 11 we see that we cannot conclude that $e^{X+Y} = e^X e^Y$ for $n \times n$ matrices X and Y . However, a more limited property is true.

- (a) Follow the steps indicated to show that if A is an $n \times n$ matrix and s and t are any scalars, then $e^{As} e^{At} = e^{A(s+t)}$. (Although we will not use it, you may assume that the series for e^A converges for any square matrix A .)

- i. Use the definition to show that

$$e^{As} e^{At} = \sum_{k \geq 0} \sum_{m \geq 0} \frac{s^k t^m}{k! m!} A^{k+m}.$$



ii. Relabel and reorder terms with $n = k + m$ to show that

$$e^{As} e^{At} = \sum_{n \geq 0} \frac{1}{n!} A^n \sum_{m=0}^n \frac{n!}{(n-m)!m!} s^{n-m} t^m.$$

iii. Complete the problem using the Binomial Theorem that says

$$(s+t)^n = \sum_{m=0}^n \frac{n!}{(n-m)!m!} s^{n-m} t^m.$$

- (b) Use the result of part (a) to show that e^A is an invertible matrix for any $n \times n$ matrix A .

- (13) There is an interesting connection between the determinant of a matrix exponential and the trace of the matrix. Let A be a diagonalizable $n \times n$ matrix with real entries. Let $D = P^{-1}AP$ for some invertible matrix P , where D is the diagonal matrix with entries $\lambda_1, \lambda_2, \dots, \lambda_n$ the eigenvalues of A .

- (a) Show that $e^A = Pe^D P^{-1}$.

- (b) Use Exercise 9 to show that

$$\det(e^A) = e^{\text{trace}(A)}.$$

- (14) Label each of the following statements as True or False. Provide justification for your response.

- (a) **True/False** If matrix A is diagonalizable, then so is A^T .
- (b) **True/False** If matrix A is diagonalizable, then A is invertible.
- (c) **True/False** If an $n \times n$ matrix A is diagonalizable, then A has n distinct eigenvalues.
- (d) **True/False** If matrix A is invertible and diagonalizable, then so is A^{-1} .
- (e) **True/False** If an $n \times n$ matrix C is diagonalizable, then there exists a basis of \mathbb{R}^n consisting of the eigenvectors of C .
- (f) **True/False** An $n \times n$ matrix with n distinct eigenvalues is diagonalizable.
- (g) **True/False** If A is an $n \times n$ diagonalizable matrix, then there is a unique diagonal matrix such that $P^{-1}AP = D$ for some invertible matrix P .
- (h) **True/False** If A is an $n \times n$ matrix with eigenvalue λ , then the dimension of the eigenspace of A corresponding to the eigenvalue λ is $n - \text{rank}(A - \lambda I_n)$.
- (i) **True/False** If λ is an eigenvalue of an $n \times n$ matrix A , then e^λ is an eigenvalue of e^A . (See Exercise 12 in Section 8 for information on the matrix exponential.)

Project: Binet's Formula for the Fibonacci Numbers

We return to the Fibonacci sequence F_n where $F_{n+2} = F_{n+1} + F_n$, for $n \geq 0$, $F_0 = 0$, and $F_1 = 1$. Since F_{n+2} is determined by previous values F_{n+1} and F_n , the relation $F_{n+2} = F_{n+1} + F_n$ is



called a *recurrence relation*. The recurrence relation $F_{n+2} = F_{n+1} + F_n$ is very time consuming to use to compute F_n for large values of n . It turns out that there is a fascinating formula that gives the n th term of the Fibonacci sequence directly, without using the relation $F_{n+2} = F_{n+1} + F_n$.

Project Activity 18.1. The recurrence relation $F_{n+2} = F_{n+1} + F_n$ gives the equations

$$F_{n+1} = F_n + F_{n-1} \quad (18.3)$$

$$F_n = F_n. \quad (18.4)$$

Let $\mathbf{x}_n = \begin{bmatrix} F_{n+1} \\ F_n \end{bmatrix}$ for $n \geq 0$. Explain how the equations (18.3) and (18.3) can be described with the matrix equation

$$\mathbf{x}_n = A\mathbf{x}_{n-1}, \quad (18.5)$$

where $A = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}$.

The matrix equation (18.5) shows us how to find the vectors \mathbf{x}_n using powers of the matrix A :

$$\begin{aligned} \mathbf{x}_1 &= A\mathbf{x}_0 \\ \mathbf{x}_2 &= A\mathbf{x}_1 = A(A\mathbf{x}_0) = A^2\mathbf{x}_0 \\ \mathbf{x}_3 &= A\mathbf{x}_2 = A(A^2\mathbf{x}_0) = A^3\mathbf{x}_0 \\ &\vdots & \vdots \\ \mathbf{x}_n &= A^n\mathbf{x}_0. \end{aligned}$$

So if we can somehow easily find the powers of the matrix A , then we can find a convenient formula for F_n . As we have seen, we know how to do this if A is diagonalizable

Project Activity 18.2. Let $A = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}$.

(a) Show that the eigenvalues of A are $\varphi = \frac{1+\sqrt{5}}{2}$ and $\bar{\varphi} = \frac{1-\sqrt{5}}{2}$.

(b) Find bases for each eigenspace of A .

Now that we have the eigenvalues and know corresponding eigenvectors for A , we can return to the problem of diagonalizing A .

Project Activity 18.3.

(a) Why do we know that A is diagonalizable?

(b) Find a matrix P such that $P^{-1}AP$ is a diagonal matrix. What is the diagonal matrix?

Now we can find a formula for the n th Fibonacci number.

Project Activity 18.4. Since $P^{-1}AP = D$, where D is a diagonal matrix, we also have $A = PDP^{-1}$. Recall that when $A = PDP^{-1}$, it follows that $A^n = PD^nP^{-1}$. Use the equation $A^n = PD^nP^{-1}$ to show that

$$F_n = \frac{\varphi^n - \bar{\varphi}^n}{\sqrt{5}}. \quad (18.6)$$

(Hint: We just need to calculate the second component of $A^n\mathbf{x}_0$.)



Formula (18.6) is called *Binet's formula*. It is a very surprising formula in the fact that the expression on the right hand side of (18.6) is an integer for each positive integer n . Note that with Binet's formula we can quickly compute F_n for very large values of n . For example,

$$F_{150} = 9969216677189303386214405760200.$$

The number $\varphi = \frac{1+\sqrt{5}}{2}$, called the *golden mean* or *golden ratio* is intimately related to the Fibonacci sequence. Binet's formula provides a fascinating relationship between the Fibonacci numbers and the golden ratio. The golden ratio also occurs often in other areas of mathematics. It was an important number to the ancient Greek mathematicians who felt that the most aesthetically pleasing rectangles had sides in the ratio of $\varphi : 1$.

Project Activity 18.5. You might wonder what happens if we use negative integer exponents in Binet's formula. In other words, are there negatively indexed Fibonacci numbers? For any integer n , including negative integers, let

$$F_n = \frac{\varphi^n - \bar{\varphi}^n}{\sqrt{5}}$$

There is a specific relationship between F_{-n} and F_n . Find it and verify it.

Section 19

Approximating Eigenvalues and Eigenvectors

Focus Questions

By the end of this section, you should be able to give precise and thorough answers to the questions listed below. You may want to keep these questions in mind to focus your thoughts as you complete the section.

- What is the power method for?
- How does the power method work?
- How can we use the inverse power method to approximate any eigenvalue/eigenvector pair?

Application: Leslie Matrices and Population Modeling

The Leslie Matrix (also called the Leslie Model) is a powerful model for describing an age distributed growth of a population that is closed to migration. In a Leslie model, it is usually the case that only one gender (most often female) is considered. As an example, we will later consider a population of sheep that is being grown commercially. A natural question that we will address is how we can harvest the population to build a sustainable environment.

When working with populations, the matrices we use are often large. For large matrices, using the characteristic polynomial to calculate eigenvalues is too time and resource consuming to be practical, and we generally cannot find the exact values of the eigenvalues. As a result, approximation techniques are very important. In this section we will explore a method for approximating eigenvalues. The eigenvalues of a Leslie matrix are important because they describe the limiting or steady-state behavior of a population. The matrix and model were introduced by Patrick H. Leslie in “On the Use of Matrices in Certain Population Mathematics”, Leslie, P.H., *Biometrika*, Volume XXXIII, November 1945, pp. 183-212.

Introduction

We have used the characteristic polynomial to find the eigenvalues of a matrix, and for each eigenvalue row reduced a corresponding matrix to find the eigenvectors. This method is only practical for small matrices – for more realistic applications approximation techniques are used. We investigate one such technique in this section, the *power method*.

Preview Activity 19.1. Let $A = \begin{bmatrix} 2 & 6 \\ 5 & 3 \end{bmatrix}$. Our goal is to find a scalar λ and a nonzero vector \mathbf{v} so that $A\mathbf{v} = \lambda\mathbf{v}$.

- (1) If we have no prior knowledge of the eigenvalues and eigenvectors of this matrix, we might just begin with a guess. Let $\mathbf{x}_0 = [1 \ 0]^\top$ be such a guess for an eigenvector. Calculate $A\mathbf{x}_0$. Is \mathbf{x}_0 an eigenvector of A ? Explain.
- (2) If \mathbf{x}_0 is not a good approximation to an eigenvector of A , then we need to make a better guess. We have little to work with other than just random guessing, but we can use $\mathbf{x}_1 = A\mathbf{x}_0$ as another guess. We calculated \mathbf{x}_1 in part 1. Is \mathbf{x}_1 an eigenvector for A ? Explain.
- (3) In parts (a) and (b) you might have noticed that in some sense \mathbf{x}_1 is closer to being an eigenvector of A than \mathbf{x}_0 was. So maybe continuing this process will get us closer to an eigenvector of A . In other words, for each positive integer k we define \mathbf{x}_k as $A\mathbf{x}_{k-1}$. Before we proceed, however, we should note that as we calculate the vectors $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \dots$, the entries in the vectors get large very quickly. So it will be useful to scale the entries so that they stay at a reasonable size, which makes it easier to interpret the output. One way to do this is to divide each vector \mathbf{x}_i by its largest component in absolute value so that all of the entries stay between -1 and 1 .¹ So in our example we have $\mathbf{x}_0 = [1 \ 0]^\top$, $\mathbf{x}_1 = [2/5 \ 1]^\top$, and $\mathbf{x}_2 = [1 \ 25/34]^\top$. Explain why scaling our vectors will not affect our search for an eigenvector.
- (4) Use an appropriate technological tool to find the vectors \mathbf{x}_k up to $k = 10$. What do you think the limiting vector $\lim_{k \rightarrow \infty} \mathbf{x}_k$ is? Is this limiting vector an eigenvector of A ? If so, what is the corresponding eigenvalue?

The Power Method

While the examples we present in this text are small in order to highlight the concepts, matrices that appear in real life applications are often enormous. For example, in Google's PageRank algorithm that is used to determine relative rankings of the importance of web pages, matrices of staggering size are used (most entries in the matrices are zero, but the size of the matrices is still huge). Finding eigenvalues of such large matrices through the characteristic polynomial is impractical. In fact, finding the roots of all but the smallest degree characteristic polynomials is a very difficult problem. As a result, using the characteristic polynomial to find eigenvalues and then finding eigenvectors is not very practical in general, and it is often a better option to use a numeric approximation method. We will consider one such method in this section, the *power method*.

¹There are several other ways to scale, but we won't consider them here.



In Preview Activity 19.1, we saw an example of a matrix $A = \begin{bmatrix} 2 & 6 \\ 5 & 3 \end{bmatrix}$ so that the sequence $\{\mathbf{x}_k\}$, where $\mathbf{x}_k = A\mathbf{x}_{k-1}$, converged to a dominant eigenvector of A for an initial guess vector $\mathbf{x}_0 = [1 \ 0]^T$. The vectors \mathbf{x}_i for i from 1 to 6 (with scaling) are approximately

$$\begin{aligned}\mathbf{x}_1 &= \begin{bmatrix} 0.4 \\ 1 \end{bmatrix} & \mathbf{x}_2 &= \begin{bmatrix} 1 \\ 0.7353 \end{bmatrix} & \mathbf{x}_3 &= \begin{bmatrix} 0.8898 \\ 1 \end{bmatrix} \\ \mathbf{x}_4 &= \begin{bmatrix} 1 \\ 0.9575 \end{bmatrix} & \mathbf{x}_5 &= \begin{bmatrix} 0.9838 \\ 1 \end{bmatrix} & \mathbf{x}_6 &= \begin{bmatrix} 0.9939 \\ 1 \end{bmatrix}.\end{aligned}$$

Numerically we can see that the sequence $\{\mathbf{x}_k\}$ approaches the vector $[1 \ 1]^T$, and Figure 19.1 illustrates this geometrically as well. This method of successive approximations $\mathbf{x}_k = A\mathbf{x}_{k-1}$ is

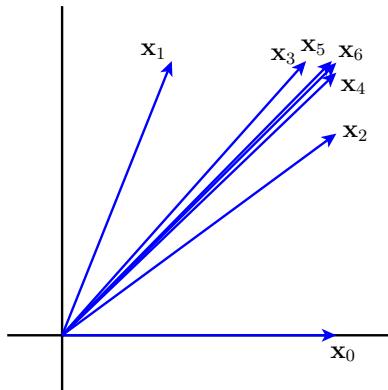


Figure 19.1: The power method.

called the *power method* (since we could write \mathbf{x}_k as $A^k \mathbf{x}_0$). Our task now is to show that this method works in general. In the next activity we restrict our argument to the 2×2 case, and then discuss the general case afterwards.

Let A be an arbitrary 2×2 matrix with two linearly independent eigenvectors \mathbf{v}_1 and \mathbf{v}_2 and corresponding eigenvalues λ_1 and λ_2 , respectively. We will also assume $|\lambda_1| > |\lambda_2|$. An eigenvalue whose absolute value is larger than that of any other eigenvalue is called a *dominant eigenvalue*. Any eigenvector for a dominant eigenvalue is called a *dominant eigenvector*. Before we show that our method can be used to approximate a dominant eigenvector, we recall that since \mathbf{v}_1 and \mathbf{v}_2 are eigenvectors corresponding to distinct eigenvalues, then \mathbf{v}_1 and \mathbf{v}_2 are linearly independent. So there exist scalars a_1 and a_2 such that

$$\mathbf{x}_0 = a_1 \mathbf{v}_1 + a_2 \mathbf{v}_2.$$

We have seen that for each positive integer k we can write \mathbf{x}_n as

$$\mathbf{x}_k = a_1 \lambda_1^k \mathbf{v}_1 + a_2 \lambda_2^k \mathbf{v}_2. \quad (19.1)$$

With this representation of \mathbf{x}_0 we can now see why the power method approximates a dominant eigenvector of A .

Activity 19.1. Assume as above that A is an arbitrary 2×2 matrix with two linearly independent eigenvectors \mathbf{v}_1 and \mathbf{v}_2 and corresponding eigenvalues λ_1 and λ_2 , respectively. (We are assuming that we don't know these eigenvectors, but we can assume that they exist.) Assume that λ_1 is the dominant eigenvalue for A , \mathbf{x}_0 is some initial guess to an eigenvector for A , that $\mathbf{x}_0 = a_1\mathbf{v}_1 + a_2\mathbf{v}_2$, and that $\mathbf{x}_k = A\mathbf{x}_{k-1}$ for $k \geq 1$.

- (a) We divide both sides of equation (19.1) by λ_1^k (since λ_1 is the dominant eigenvalue, we know that λ_1 is not 0) to obtain

$$\frac{1}{\lambda_1^k} \mathbf{x}_k = a_1 \mathbf{v}_1 + a_2 \left(\frac{\lambda_2}{\lambda_1} \right)^k \mathbf{v}_2. \quad (19.2)$$

Recall that λ_1 is the dominant eigenvalue for A . What happens to $\left(\frac{\lambda_2}{\lambda_1} \right)^k$ as $k \rightarrow \infty$? Explain what happens to the right hand side of equation (19.2) as $k \rightarrow \infty$.

- (b) Explain why the previous result tells us that the vectors \mathbf{x}_k are approaching a vector in the *direction* of \mathbf{v}_1 or $-\mathbf{v}_1$ as $k \rightarrow \infty$, assuming $a_1 \neq 0$. (Why do we need $a_1 \neq 0$? What happens if $a_1 = 0$?)
- (c) What does all of this tell us about the sequence $\{\mathbf{x}_k\}$ as $k \rightarrow \infty$?

The power method is straightforward to implement, but it is not without its drawbacks. We began by assuming that we had a basis of eigenvectors of a matrix A . So we are also assuming that A is diagonalizable. We also assumed that A had a dominant eigenvalue λ_1 . That is, if A is $n \times n$ we assume that A has eigenvalues $\lambda_1, \lambda_2, \dots, \lambda_n$, not necessarily distinct, with

$$|\lambda_1| > |\lambda_2| \geq |\lambda_3| \geq \cdots \geq |\lambda_n|$$

and with \mathbf{v}_i an eigenvector of A with eigenvalue λ_i . We could then write any initial guess \mathbf{x}_0 in the form

$$\mathbf{x}_0 = a_1 \mathbf{v}_1 + a_2 \mathbf{v}_2 + \cdots + a_n \mathbf{v}_n.$$

The initial guess is also called a *seed*.

Then

$$\mathbf{x}_k = a_1 \lambda_1^k \mathbf{v}_1 + a_2 \lambda_2^k \mathbf{v}_2 + \cdots + a_n \lambda_n^k \mathbf{v}_n$$

and

$$\frac{1}{\lambda_1^k} \mathbf{x}_k = a_1 \mathbf{v}_1 + a_2 \left(\frac{\lambda_2}{\lambda_1} \right)^k \mathbf{v}_2 + \cdots + a_n \left(\frac{\lambda_n}{\lambda_1} \right)^k \mathbf{v}_n. \quad (19.3)$$

Notice that we are not actually calculating the vectors \mathbf{x}_k here – this is a theoretical argument and we don't know λ_1 and are not performing any scaling like we did in Preview Activity 19.1. We are assuming that λ_1 is the dominant eigenvalue of A , though, so for each i the terms $\left(\frac{\lambda_i}{\lambda_1} \right)^k$ converge to 0 as k goes to infinity. Thus,

$$\mathbf{x}_k \approx \lambda_1^k a_1 \mathbf{v}_1$$

for large values of k , which makes the sequence $\{\mathbf{x}_k\}$ converge to a vector in the direction of a dominant eigenvector \mathbf{v}_1 provided $a_1 \neq 0$. So we need to be careful enough to choose a seed that has a nonzero component in the direction of \mathbf{v}_1 . Of course, we generally don't know that our matrix



is diagonalizable before we make these calculations, but for many matrices the sequence $\{\mathbf{x}_k\}$ will approach a dominant eigenvector.

Once we have an approximation to a dominant eigenvector, we can then approximate the dominant eigenvalue.

Activity 19.2. Let A be an $n \times n$ matrix with eigenvalue λ and corresponding eigenvector \mathbf{v} .

(a) Explain why $\lambda = \frac{\lambda(\mathbf{v} \cdot \mathbf{v})}{\mathbf{v} \cdot \mathbf{v}}$.

(b) Use the result of part (a) to explain why $\lambda = \frac{(A\mathbf{v}) \cdot \mathbf{v}}{\mathbf{v} \cdot \mathbf{v}}$.

The result of Activity 19.2 is that, when the vectors in the sequence $\{\mathbf{x}_k\}$ approximate a dominant eigenvector of a matrix A , the quotients

$$\frac{(A\mathbf{x}_k) \cdot \mathbf{x}_k}{\mathbf{x}_k \cdot \mathbf{x}_k} = \frac{\mathbf{x}_k^T A \mathbf{x}_k}{\mathbf{x}_k^T \mathbf{x}_k} \quad (19.4)$$

approximate the dominant eigenvalue of A . The quotients in (19.4) are called *Rayleigh quotients*.

To summarize, the procedure for applying the power method for approximating a dominant eigenvector and dominant eigenvalue of a matrix A is as follows.

Step 1: Select an arbitrary nonzero vector \mathbf{x}_0 as an initial guess to a dominant eigenvector.

Step 2: Let $\mathbf{x}_1 = A\mathbf{x}_0$. Let $k = 1$.

Step 3: To avoid having the magnitudes of successive approximations become excessively large, scale this approximation \mathbf{x}_k . That is, find the entry α_k of \mathbf{x}_k that is largest in absolute value. Then replace \mathbf{x}_k by $\frac{1}{|\alpha_k|}\mathbf{x}_k$. Note that this does not change the direction of this approximation, only its magnitude.

Step 4: Calculate the Rayleigh quotient $r_k = \frac{(A\mathbf{x}_k) \cdot \mathbf{x}_k}{\mathbf{x}_k \cdot \mathbf{x}_k}$.

Step 5: Let let $\mathbf{x}_{k+1} = A\mathbf{x}_k$. Increase k by 1 and repeat Steps 3 through 5.

If the sequence $\{\mathbf{x}_k\}$ converges to a dominant eigenvector of A , then the sequence $\{r_k\}$ converges to the dominant eigenvalue of A .

The power method can be useful for approximating a dominant eigenvector as long as the successive multiplications by A are fairly simple – for example, if many entries of A are zero.² The rate of convergence of the sequence $\{\mathbf{x}_k\}$ depends on the ratio $\frac{\lambda_2}{\lambda_1}$. If this ratio is close to 1, then it can take many iterations before the power $\left(\frac{\lambda_2}{\lambda_1}\right)^k$ makes the \mathbf{v}_2 term negligible. There are other methods for approximating eigenvalues and eigenvectors, e.g., the QR factorization, that we will not discuss at this point.

²A matrix in which most entries are zero is called a *sparse* matrix.



The Inverse Power Method

The power method only allows us to approximate the dominant eigenvalue and a dominant eigenvector for a matrix A . It is possible to modify this method to approximate other eigenvectors and eigenvalues under certain conditions. We consider an example in the next activity to motivate the general situation.

Activity 19.3. Let $A = \begin{bmatrix} 2 & 6 \\ 5 & 3 \end{bmatrix}$ be the matrix from Preview Activity 19.1. Recall that 8 is an eigenvalue for A , and a quick calculation can show that -3 is the other eigenvalue of A . Consider the matrix $B = (A - (-2)I_2)^{-1} = \frac{1}{10} \begin{bmatrix} -5 & 6 \\ 5 & -4 \end{bmatrix}$.

- (a) Show that $\frac{1}{8-(-2)}$ and $\frac{1}{-3-(-2)}$ are the eigenvalues of B .
- (b) Recall that $\mathbf{v}_1 = [1 \ 1]^T$ is an eigenvector of A corresponding to the eigenvalue 8 and assume that $\mathbf{v}_2 = [-6 \ 5]^T$ is an eigenvector for A corresponding to the eigenvalue -3 . Calculate the products $B\mathbf{v}_1$ and $B\mathbf{v}_2$. How do the products relate to the results of part (a)?

Activity 19.3 provides evidence that we can translate the matrix A having a dominant eigenvalue to a different matrix B with the same eigenvectors as A and with a dominant eigenvalue of our choosing. To see why, let A be an $n \times n$ matrix with eigenvalues $\lambda_1, \lambda_2, \dots, \lambda_n$, and let α be any real number distinct from the eigenvalues. Let $B = (A - \alpha I_n)^{-1}$. In our example in Activity 19.3 the numbers

$$\frac{1}{\lambda_1 - \alpha}, \frac{1}{\lambda_2 - \alpha}, \frac{1}{\lambda_3 - \alpha}, \dots, \frac{1}{\lambda_n - \alpha}$$

were the eigenvalues of B , and that if \mathbf{v}_i is an eigenvector for A corresponding to the eigenvalue λ_i , then \mathbf{v}_i is an eigenvector of B corresponding to the eigenvalue $\frac{1}{\lambda_i - \alpha}$. To see why, let λ be an eigenvalue of an $n \times n$ matrix A with corresponding eigenvector \mathbf{v} . Let α be a scalar that is not an eigenvalue of A , and let $B = (A - \alpha I_n)^{-1}$. Now

$$\begin{aligned} A\mathbf{v} &= \lambda\mathbf{v} \\ A\mathbf{v} - \alpha\mathbf{v} &= \lambda\mathbf{v} - \alpha\mathbf{v} \\ (A - \alpha I_n)\mathbf{v} &= (\lambda - \alpha)\mathbf{v} \\ \frac{1}{\lambda - \alpha}\mathbf{v} &= (A - \alpha I_n)^{-1}\mathbf{v}. \end{aligned}$$

So $\frac{1}{\lambda - \alpha}$ is an eigenvalue of B with eigenvector \mathbf{v} .

Now suppose that A is an $n \times n$ matrix with eigenvalues $\lambda_1, \lambda_2, \dots, \lambda_n$, and that we want to approximate an eigenvector and corresponding eigenvalue λ_i of A . If we can somehow find a value of α so that $|\lambda_i - \alpha| < |\lambda_j - \alpha|$ for all $j \neq i$, then $\left| \frac{1}{\lambda_i - \alpha} \right| > \left| \frac{1}{\lambda_j - \alpha} \right|$ for any $j \neq i$. Thus, the matrix $B = (A - \alpha I_n)^{-1}$ has $\frac{1}{\lambda_i - \alpha}$ as its dominant eigenvalue and we can use the power method to approximate an eigenvector and the Rayleigh quotient to approximate the eigenvalue $\frac{1}{\lambda_i - \alpha}$, and hence approximate λ_i .

Activity 19.4. Let $A = \frac{1}{8} \begin{bmatrix} 7 & 3 & 3 \\ 30 & 22 & -10 \\ 15 & -21 & 11 \end{bmatrix}$.



- (a) Apply the power method to the matrix $B = (A - I_3)^{-1}$ with initial vector $\mathbf{x}_0 = [1 \ 0 \ 0]^\top$ to fill in Table 19.1 (to four decimal places). Use this information to estimate an eigenvalue for A and a corresponding eigenvector.

k	10	15	20
\mathbf{x}_k			
$\frac{\mathbf{x}_k^\top A \mathbf{x}_k}{\mathbf{x}_k^\top \mathbf{x}_k}$			

Table 19.1: Applying the power method to $(A - I_3)^{-1}$.

- (b) Applying the power method to the matrix $B = (A - 0I_3)^{-1}$ with initial vector $\mathbf{x}_0 = [1 \ 0 \ 0]^\top$ yields the information in Table 19.2 (to four decimal places). Use this information to estimate an eigenvalue for A and a corresponding eigenvector.

k	10	15	20
\mathbf{x}_k	$\begin{bmatrix} 0.3344 \\ -0.6677 \\ -1.0000 \end{bmatrix}$	$\begin{bmatrix} -0.3333 \\ 0.6666 \\ 1.0000 \end{bmatrix}$	$\begin{bmatrix} 0.3333 \\ -0.6666 \\ -1.0000 \end{bmatrix}$
$\frac{\mathbf{x}_k^\top A \mathbf{x}_k}{\mathbf{x}_k^\top \mathbf{x}_k}$	-1.0014	-1.0000	-1.0000

Table 19.2: Applying the power method to $(A - 0I_3)^{-1}$.

- (c) Applying the power method to the matrix $B = (A - 5I_3)^{-1}$ with initial vector $\mathbf{x}_0 = [1 \ 0 \ 0]^\top$ yields the information in Table 19.3 (to four decimal places). Use this information to estimate an eigenvalue for A and a corresponding eigenvector.

Examples

What follows are worked examples that use the concepts from this section.

Example 19.1. Let $A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}$.

- (a) Approximate the dominant eigenvalue of A accurate to two decimal places using the power method. Use technology as appropriate.



k	10	15	20
\mathbf{x}_k	$\begin{bmatrix} 0.0000 \\ 1.0000 \\ -1.0000 \end{bmatrix}$	$\begin{bmatrix} 0.0000 \\ -1.0000 \\ 1.0000 \end{bmatrix}$	$\begin{bmatrix} 0.0000 \\ 1.0000 \\ -1.0000 \end{bmatrix}$
$\frac{\mathbf{x}_k^T A \mathbf{x}_k}{\mathbf{x}_k^T \mathbf{x}_k}$	-1.0000	-1.0000	-1.0000

Table 19.3: Applying the power method to $(A - 5I_3)^{-1}$.

- (b) Find the characteristic polynomial $p(\lambda)$ of A . Then find the root of $p(\lambda)$ farthest from the origin. Compare to the result of part (a). Use technology as appropriate.

Example Solution.

- (a) We use technology to calculate the scaled vectors $A^k \mathbf{x}_0$ for values of k until the components don't change in the second decimal place. We start with the seed $\mathbf{x}_0 = [1 \ 1 \ 1]^T$. For example, to two decimal places we have $\mathbf{x}_k = [0.28 \ 0.64 \ 1.00]^T$ for $k \geq 20$. So we suspect that $[0.28 \ 0.64 \ 1.00]^T$ is close to a dominant eigenvector for A .

For the dominant eigenvalue, we can calculate the Rayleigh quotients $\frac{(A\mathbf{x}_k) \cdot \mathbf{x}_k}{\mathbf{x}_k^T \mathbf{x}_k}$ until they do not change to two decimal places. For $k \geq 4$, our Rayleigh quotients are all (to two decimal places) equal to 16.12. So we expect that the dominant eigenvalue of A is close to 16.12. Notice that

$$A[0.28 \ 0.64 \ 1.00]^T = [4.56 \ 10.32 \ 16.08]^T,$$

which is not far off from $16.12[0.28 \ 0.64 \ 1.00]^T$.

- (b) The characteristic polynomial of A is

$$p(\lambda) = -\lambda^3 + 15\lambda^2 + 18\lambda = -\lambda(\lambda^2 - 15\lambda - 18).$$

The quadratic formula gives the nonzero roots of $p(\lambda)$ as

$$\frac{15 \pm \sqrt{15^2 + 4(18)}}{2} = \frac{15 \pm 3\sqrt{33}}{2}.$$

The roots farthest from the origin is approximately 16.12, as was also calculated in part (a).

Example 19.2. Let $A = \begin{bmatrix} 2 & 1 & 0 \\ 1 & 3 & 1 \\ 0 & 1 & 2 \end{bmatrix}$.

- (a) Use the power method to approximate the dominant eigenvalue and a corresponding eigenvector (using scaling) accurate to two decimal places. Use $\mathbf{x}_0 = [1 \ 1 \ 1]^T$ as the seed.



- (b) Determine the exact value of the dominant eigenvalue of A and compare to your result from part (a).
- (c) Approximate the remaining eigenvalues of A using the inverse power method. (Hint: Try $\alpha = 0.5$ and $\alpha = 1.8$.)

Example Solution.

- (a) We use technology to calculate the scaled vectors $A^k \mathbf{x}_0$ for values of k until the components don't change in the second decimal place. For example, to two decimal places we have $\mathbf{x}_k = [0.50 \ 1.00 \ 0.50]^\top$ for $k \geq 4$. So we suspect that $\left[\frac{1}{2} \ 1 \ \frac{1}{2}\right]^\top$ is a dominant eigenvector for A .

For the dominant eigenvalue, we can calculate the Rayleigh quotients $\frac{(A\mathbf{x}_k) \cdot \mathbf{x}_k}{\mathbf{x}_k \cdot \mathbf{x}_k}$ until they do not change to two decimal places. For $k \geq 2$, our Rayleigh quotients are all (to two decimal places) equal to 4. So we expect that the dominant eigenvalue of A is 4. We could also use the fact that

$$A \begin{bmatrix} \frac{1}{2} & 1 & \frac{1}{2} \end{bmatrix}^\top = [2 \ 4 \ 2]^\top = 4 \begin{bmatrix} \frac{1}{2} & 1 & \frac{1}{2} \end{bmatrix}^\top$$

to see that $\left[\frac{1}{2} \ 1 \ \frac{1}{2}\right]^\top$ is a dominant eigenvector for A with eigenvalue 4.

- (b) Technology shows that the characteristic polynomial of $A - \lambda I_3$ is

$$p(\lambda) = -\lambda^3 + 7\lambda^2 - 14\lambda + 8 = -(\lambda - 1)(\lambda - 2)(\lambda - 4).$$

We can see from the characteristic polynomial that 4 is the dominant eigenvalue of A .

- (c) Applying the power method to $B = (A - 0.5I_3)^{-1}$ with seed $\mathbf{x}_0 = [1 \ 1 \ 1]^\top$ gives $\mathbf{x}_k \approx [0.50 \ 1.00 \ 0.50]^\top$ for $k \geq 5$, with Rayleigh quotients of 2 (to several decimal places). So 2 is the dominant eigenvalue of B . But $\frac{1}{\lambda-0.5}$ is also the dominant eigenvalue of B , where λ is the corresponding eigenvalue of A . So to find λ , we note that $\frac{1}{\lambda-0.5} = 2$ implies that $\lambda = 1$ is an eigenvalue of A .

Now applying the power method to $B = (A - 1.8I_3)^{-1}$ with seed $\mathbf{x}_0 = [1 \ 1 \ 1]^\top$ gives $\mathbf{x}_k \approx [1.00 \ -1.00 \ 1.00]^\top$ for large enough k , with Rayleigh quotients of 5 (to several decimal places). To find the corresponding eigenvalue λ for A , we note that $\frac{1}{\lambda-1.8} = 5$, or $\lambda = 2$ is an eigenvalue of A .

Admittedly, this method is very limited. Finding good choices for α often depends on having some information about the eigenvalues of A . Choosing α close to an eigenvalue provides the best chance of obtaining that eigenvalue.

Summary

- The power method is an iterative method that can be used to approximate the dominant eigenvalue of an $n \times n$ matrix A that has n linearly independent eigenvectors and a dominant eigenvalue.



- To use the power method we start with a seed \mathbf{x}_0 and then calculate the sequence $\{\mathbf{x}_k\}$ of vectors, where $\mathbf{x}_k = A\mathbf{x}_{k-1}$. If \mathbf{x}_0 is chosen well, then the sequence $\{\mathbf{x}_k\}$ converges to a dominant eigenvector of A .
- If A is an $n \times n$ matrix with eigenvalues $\lambda_1, \lambda_2, \dots, \lambda_n$, to approximate an eigenvector of A corresponding to the eigenvalue λ_i , we apply the power method to the matrix $B = (A - \alpha I_n)^{-1}$, where α is not a eigenvalue of A and $\left| \frac{1}{\lambda_i - \alpha} \right| > \left| \frac{1}{\lambda_j - \alpha} \right|$ for any $j \neq i$.

Exercises

(1) Let $A = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}$. Let $\mathbf{x}_0 = [1 \ 0]^T$.

- Find the eigenvalues and corresponding eigenvectors for A .
- Use appropriate technology to calculate $\mathbf{x}_k = A^k \mathbf{x}_0$ for k up to 10. Compare to a dominant eigenvector for A .
- Use the eigenvectors from part (b) to approximate the dominant eigenvalue for A . Compare to the exact value of the dominant eigenvalue of A .
- Assume that the other eigenvalue for A is close to 0. Apply the inverse power method and compare the results to the remaining eigenvalue and eigenvectors for A .

(2) Let $A = \begin{bmatrix} 1 & 2 & 0 \\ -2 & 1 & 2 \\ 1 & 3 & 1 \end{bmatrix}$. Use the power method to approximate a dominant eigenvector for A . Use $\mathbf{x}_0 = [1 \ 1 \ 1]^T$ as the seed. Then approximate the dominant eigenvalue of A .

(3) Let $A = \begin{bmatrix} 3 & -1 \\ -1 & 3 \end{bmatrix}$. Use the power method starting with $\mathbf{x}_0 = [1 \ 1]^T$. Explain why the method fails in this case to approximate a dominant eigenvector, and how you could adjust the seed to make the process work.

(4) Let $A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$.

- Find the eigenvalues and an eigenvector for each eigenvalue.
- Apply the power method with an initial starting vector $\mathbf{x}_0 = [0 \ 1]^T$. What is the resulting sequence?
- Use equation (19.3) to explain the sequence you found in part (b).

(5) Let $A = \begin{bmatrix} 2 & 6 \\ 5 & 3 \end{bmatrix}$. Fill in the entries in Table 19.4, where \mathbf{x}_k is the k th approximation to a dominant eigenvector using the power method, starting with the seed $\mathbf{x}_0 = [1 \ 0]^T$. Compare the results of this table to the eigenvalues of A and $\lim_{k \rightarrow \infty} \frac{\mathbf{x}_{k+1} \cdot \mathbf{x}_k}{\mathbf{x}_k \cdot \mathbf{x}_k}$. What do you notice?

(6) Let $A = \begin{bmatrix} 4 & -5 \\ 2 & 15 \end{bmatrix}$. The power method will approximate the dominant eigenvalue $\lambda = 14$. In this exercise we explore what happens if we apply the power method to A^{-1} .

\mathbf{v}	\mathbf{x}_0	\mathbf{x}_1	\mathbf{x}_2	\mathbf{x}_3	\mathbf{x}_4	\mathbf{x}_5
$\frac{\mathbf{v}^T A \mathbf{v}}{\mathbf{v}^T \mathbf{v}}$						
\mathbf{v}	\mathbf{x}_6	\mathbf{x}_7	\mathbf{x}_8	\mathbf{x}_9	\mathbf{x}_{10}	\mathbf{x}_{11}
$\frac{\mathbf{v}^T A \mathbf{v}}{\mathbf{v}^T \mathbf{v}}$						

Table 19.4: Values of the Rayleigh quotient.

- (a) Apply the power method to A^{-1} to approximate the dominant eigenvalue of A^{-1} . Use $[1 \ 1]^T$ as the seed. How is this eigenvalue related to an eigenvalue of A ?
- (b) Explain in general why applying the power method to the inverse of an invertible matrix B might give an approximation to an eigenvalue of B of smallest magnitude. When might this not work?
- (7) There are other algebraic methods that do not rely on the determinant of a matrix that can be used to find eigenvalues of a matrix. We examine one such method in this exercise. Let A be any $n \times n$ matrix, and let \mathbf{v} be any vector in \mathbb{R}^n .

- (a) Explain why the vectors

$$\mathbf{v}, A\mathbf{v}, A^2\mathbf{v}, \dots, A^n\mathbf{v}$$

are linearly independent.

- (b) Let c_0, c_1, \dots, c_n be scalars, not all 0, so that

$$c_0\mathbf{v} + c_1A\mathbf{v} + c_2A^2\mathbf{v} + \dots + c_nA^n\mathbf{v} = \mathbf{0}.$$

Explain why there must be a smallest positive integer k so that there are scalars a_0, a_1, \dots, a_k with $a_k \neq 0$, such that

$$a_0\mathbf{v} + a_1A\mathbf{v} + a_2A^2\mathbf{v} + \dots + a_kA^k\mathbf{v} = \mathbf{0}.$$

- (c) Let

$$q(t) = a_0 + a_1t + a_2t^2 + \dots + a_kt^k.$$

Then

$$q(A) = a_0 + a_1A + a_2A^2 + \dots + a_kA^k$$

and

$$\begin{aligned} q(A)\mathbf{v} &= (a_0 + a_1A + a_2A^2 + \dots + a_kA^k)\mathbf{v} \\ &= a_0\mathbf{v} + a_1A\mathbf{v} + a_2A^2\mathbf{v} + \dots + a_kA^k\mathbf{v} \\ &= \mathbf{0}. \end{aligned}$$

Suppose the polynomial $q(t)$ has a linear factor, say $q(t) = (t - \lambda)Q(t)$ for some degree $k-1$ polynomial $Q(t)$. Explain why, if $Q(A)\mathbf{v}$ is non-zero, λ is an eigenvalue of A with eigenvector $Q(A)\mathbf{v}$.



- (d) This method allows us to find certain eigenvalues and eigenvectors, the roots of the polynomial $q(t)$. Any other eigenvector must lie outside the eigenspaces we have already found, so repeating the process with a vector \mathbf{v} not in any of the known eigenspaces will produce different eigenvalues and eigenvectors. Let $A = \begin{bmatrix} 2 & 2 & -1 \\ 2 & 2 & 2 \\ 0 & 0 & 6 \end{bmatrix}$.
- Find the polynomial $q(t)$. Use $\mathbf{v} = [1 \ 1 \ 1]^T$.
 - Find all of the roots of $q(t)$.
 - For each root λ of $q(t)$, find the polynomial $Q(t)$ and use this polynomial to determine an eigenvector of A . Verify your work.
- (8) We don't need to use the Rayleigh quotients to approximate the dominant eigenvalue of a matrix A if we instead keep track of the scaling factors. Recall that the scaling in the power method can be used to make the magnitudes of the successive approximations smaller and easier to work with. Let A be an $n \times n$ matrix and begin with a non-zero seed \mathbf{v}_0 . We now want to keep track of the scaling factors, so let α_0 be the component of \mathbf{v}_0 with largest absolute value and let $\mathbf{x}_0 = \frac{1}{|\alpha_0|}\mathbf{v}_0$. For $k \geq 0$, let $\mathbf{v}_k = A\mathbf{x}_{k-1}$, let α_k be the component of \mathbf{v}_k with largest absolute value and let $\mathbf{x}_k = \frac{1}{\alpha_k}\mathbf{v}_k$.
- Let $A = \begin{bmatrix} 0 & 1 \\ -8 & 6 \end{bmatrix}$. Use $\mathbf{x}_0 = [1 \ 1]^T$ as the seed and calculate α_k for k from 1 to 10. Compare to the dominant eigenvalue of A .
 - Assume that for large k the vectors \mathbf{x}_k approach a dominant eigenvector with dominant eigenvalue λ . Show now in general that the sequence of scaling factors α_k approaches λ .
- (9) Let A be an $n \times n$ matrix and let α be a scalar that is not an eigenvalue of A . Suppose that \mathbf{x} is an eigenvector of $B = (A - \alpha I_n)^{-1}$ with eigenvalue β . Find an eigenvalue of A in terms of β and α with corresponding eigenvector \mathbf{x} .
- (10) Label each of the following statements as True or False. Provide justification for your response.
- True/False** The largest eigenvalue of a matrix is a dominant eigenvalue.
 - True/False** If an $n \times n$ matrix A has n linearly independent eigenvectors and a dominant eigenvalue, then the sequence $\{A^k \mathbf{x}_0\}$ converges to a dominant eigenvector of A for any initial vector \mathbf{x}_0 .
 - True/False** If λ is an eigenvalue of an $n \times n$ matrix A and α is not an eigenvalue of A , then $\lambda - \alpha$ is an eigenvalue of $A - \alpha I_n$.
 - True/False** Every square matrix has a dominant eigenvalue.

Project: Managing a Sheep Herd

Sheep farming is a significant industry in New Zealand. New Zealand is reported to have the highest density of sheep in the world. Sheep can begin to reproduce after one year, and give birth



only once per year. Table 19.5 gives Birth and Survival Rates for Female New Zealand Sheep (from G. Caughley, “Parameters for Seasonally Breeding Populations,” *Ecology*, **48**, (1967), 834-839). Since sheep hardly ever live past 12 years, we will only consider the population through 12 years.

Age (years)	Birth Rate	Survival Rate
0-1	0.000	0.845
1-2	0.045	0.975
2-3	0.391	0.965
3-4	0.472	0.950
4-5	0.484	0.926
5-6	0.546	0.895
6-7	0.543	0.850
7-8	0.502	0.786
8-9	0.468	0.691
9-10	0.459	0.561
10-11	0.433	0.370
11-12	0.421	0.000

Table 19.5: New Zealand female sheep data by age group.

As sheep reproduce, they add to the 0-1 sheep (lamb) population. The potential to produce offspring is called *fecundity* (derived from the word *fecund* which generally refers to reproductive ability) and determines how many lamb are added to the population. Let F_k (the fecundity rate) be the rate at which females in age class k give birth to female offspring. Not all members of a given age group survive to the next age groups, so let s_k be the fraction of individuals that survives from age group k to age group $k + 1$. With these ideas in mind, we can create a life cycle chart as in Figure 19.2 that illustrates how the population of sheep changes on a farm (for the sake of space, we illustrate with four age classes).

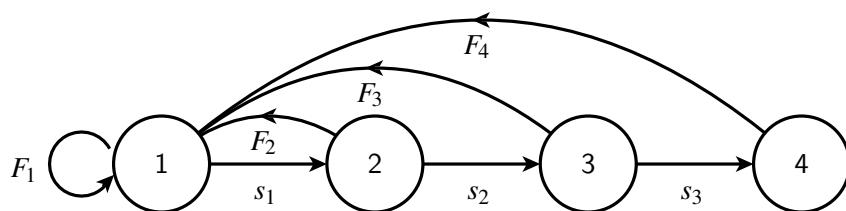


Figure 19.2: Life cycle with four age classes.

To model the sheep population, we need a few variables. Let $n_1^{(0)}$ be the number of sheep in age group 0-1, $n_2^{(0)}$ the number in age group 1-2, $n_3^{(0)}$ the number in age group 2-3 and, in general, $n_k^{(0)}$ the number of sheep in age group $(k - 1)-k$ at some initial time (time 0), and let

$$\mathbf{x}_0 = \begin{bmatrix} n_1^{(0)} & n_2^{(0)} & n_3^{(0)} & \cdots & n_{12}^{(0)} \end{bmatrix}^T.$$

We wish to determine the populations in the different groups after one year. Let

$$\mathbf{x}_1 = \begin{bmatrix} n_1^{(1)} & n_2^{(1)} & n_3^{(1)} & \cdots & n_{12}^{(1)} \end{bmatrix}^T,$$

where $n_1^{(1)}$ denotes the number of sheep in age group 0-1, $n_2^{(1)}$ the number of sheep in age group 1-2 and, in general, $n_k^{(1)}$ the number of tilapia in age group $(k - 1)$ - k after one year.

Project Activity 19.1. Table 19.5 shows that, on average, each female in age group 1-2 produces 0.045 female offspring in a year. Since there are n_2 females in age group 1-2, the lamb population increases by $0.045n_2$ in a year.

- (a) Continue this analysis to explain why

$$\begin{aligned} n_1^{(1)} = & 0.045n_2 + 0.391n_3 + 0.472n_4 + 0.484n_5 + 0.546n_6 + 0.543n_7 \\ & + 0.502n_8 + 0.468n_9 + 0.459n_{10} + 0.433n_{11} + 0.421n_{12}. \end{aligned}$$

- (b) Explain why $n_2^{(1)} = 0.845n_1$.

- (c) Now explain why

$$\mathbf{x}_1 = L\mathbf{x}_0, \quad (19.5)$$

where L is the matrix

$$\left[\begin{array}{cccccccccccc} 0 & 0.045 & 0.391 & 0.472 & 0.484 & 0.546 & 0.543 & 0.502 & 0.468 & 0.459 & 0.433 & 0.421 \\ 0.845 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0.975 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.965 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.950 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0.926 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0.895 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0.850 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.786 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.691 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.561 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.370 & 0 \end{array} \right]. \quad (19.6)$$

Notice that our matrix L has the form

$$\left[\begin{array}{cccccc} F_1 & F_2 & F_3 & \cdots & F_{n-1} & F_n \\ s_1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & s_2 & 0 & \cdots & 0 & 0 \\ 0 & 0 & s_3 & \cdots & 0 & 0 \\ & & & \ddots & & \\ 0 & 0 & 0 & \cdots & s_{n-1} & 0 \end{array} \right].$$

Such a matrix is called a *Leslie matrix*.

Leslie matrices have certain useful properties, and one eigenvalue of a Leslie matrix can tell us a lot about the long-term behavior of the situation being modeled. You can take these properties as fact unless otherwise directed.

- (1) A Leslie matrix L has a unique positive eigenvalue λ_1 with a corresponding eigenvector \mathbf{v}_1 whose entries are all positive.



- (2) If λ_i ($i > 1$) is any other eigenvalue (real or complex) of L , then $|\lambda_i| \leq \lambda_1$. If λ_1 is the largest magnitude eigenvalue of a matrix L , we call λ_1 a *dominant eigenvalue* of L .
- (3) If any two successive entries in the first row of L are both positive, then $|\lambda_i| < \lambda_1$ for every $i > 1$. In this case we say that λ_1 is a *strictly dominant eigenvalue* of L . In a Leslie model, this happens when the females in two successive age classes are fertile, which is almost always the case.
- (4) If λ_1 is a strictly dominant eigenvalue, then \mathbf{x}_k is approximately a scalar multiple of \mathbf{v}_1 for large values of k , regardless of the initial state \mathbf{x}_0 . In other words, large state vectors are close to eigenvectors for λ_1 .

We can use these properties to determine the long-term behavior of the sheep herd.

Project Activity 19.2. Assume that L is defined by (19.6), and let

$$\mathbf{x}_m = \begin{bmatrix} n_1^{(m)} & n_2^{(m)} & n_3^{(m)} & \cdots & n_{12}^{(m)} \end{bmatrix}^T,$$

where $n_1^{(m)}$ denotes the number of sheep in age group 0-1, $n_2^{(m)}$ the number of sheep in age group 1-2 and, in general, $n_k^{(m)}$ the number of sheep in age group $(k - 1)$ - k after k years.

- (a) Assume that $\mathbf{x}_0 = [100 \ 100 \ 100 \ \cdots \ 100]^T$. Use appropriate technology to calculate \mathbf{x}_{22} , \mathbf{x}_{23} , \mathbf{x}_{24} , and \mathbf{x}_{25} . Round to the nearest whole number. What do you notice about the sheep population? You may use the GeoGebra applet at <https://www.geogebra.org/m/yqss88xq>.
- (b) We can use the third and fourth properties of Leslie matrices to better understand the long-term behavior of the sheep population. Since successive entries in the first row of the Leslie matrix in (19.6) are positive, our Leslie matrix has a strictly dominant eigenvalue λ_1 . Given the dimensions of our Leslie matrix, finding this dominant eigenvalue through algebraic means is not feasible. Use the power method to approximate the dominant eigenvalue λ_1 of the Leslie matrix in (19.6) to five decimal places. Explain your process. Then explain how this dominant eigenvalue tells us that, unchecked, the sheep population grows at a rate that is roughly exponential. What is the growth rate of this exponential growth? You may use the GeoGebra applet at <https://www.geogebra.org/m/yqss88xq>.

Project Activity 19.2 indicates that, unchecked, the sheep population will grow without bound, roughly exponentially with ratio equal to the dominant eigenvalue of our Leslie matrix L . Of course, a sheep farmer cannot provide the physical environment or the resources to support an unlimited population of sheep. In addition, most sheep farmers cannot support themselves only by shearing sheep for the wool. Consequently, some harvesting of the sheep population each year for meat and skin is necessary. A sustainable harvesting policy allows for the regular harvesting of some sheep while maintaining the population at a stable level. It is necessary for the farmer to find an optimal harvesting rate to attain this stable population and the following activity leads us through an analysis of how such a harvesting rate can be determined.

Project Activity 19.3. The Leslie model can be modified to consider harvesting. It is possible to harvest different age groups at different rates, and to harvest only some age groups and not others.



In the case of sheep, it might make sense to only harvest from the youngest population since lamb is more desirable than mutton and the lamb population grows the fastest. Assume that this is our harvesting strategy and that we harvest our sheep from only the youngest age group at the start of each year. Let h be the fraction of sheep we harvest from the youngest age group each year after considering growth.

- (a) If we begin with an initial population \mathbf{x}_0 , then the state vector after births and expected deaths is $L\mathbf{x}_0$. Now we harvest. Explain why if we harvest a fraction h from the youngest age group after considering growth, then the state vector after 1 year will be

$$\mathbf{x}_1 = L\mathbf{x}_0 - H L \mathbf{x}_0,$$

where

$$H = \begin{bmatrix} h & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

- (b) Our goal is to find a harvesting rate that will lead to a steady state in which the sheep population remains the same each year. In other words, we want to find a value of h , if one exists, that satisfies

$$\mathbf{x} = L\mathbf{x} - H L \mathbf{x}. \quad (19.7)$$

Show that (19.7) is equivalent to the matrix equation

$$\mathbf{x} = (I_{12} - H)L\mathbf{x}. \quad (19.8)$$

- (c) Use appropriate technology to experiment numerically with different values of h to find the value you think gives the best uniform harvest rate. Explain your reasoning. You may use the GeoGebra applet at <https://www.geogebra.org/m/yqss88xq>.
- (d) Now we will use some algebra to find an equation that explicitly gives us the harvest rate in the general setting. This will take a bit of work, but none of it is too difficult. To simplify our work but yet illustrate the overall idea, let us consider the general 4×4 case with arbitrary Leslie matrix

$$L = \begin{bmatrix} F_1 & F_2 & F_3 & F_4 \\ s_1 & 0 & 0 & 0 \\ 0 & s_2 & 0 & 0 \\ 0 & 0 & s_3 & 0 \end{bmatrix}.$$



Recall that we want to find a value of h that satisfies (19.8) with $H = \begin{bmatrix} h & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$.

Let $\mathbf{x} = [x_1 \ x_2 \ x_3 \ x_4]^T$.

- Calculate the matrix product $(I_4 - H)L$. Explain why this product is again a Leslie matrix and why $(I_4 - H)L$ will have a dominant eigenvalue of 1.
- Now calculate $(I_4 - H)L\mathbf{x}$ and set it equal to \mathbf{x} . Write down the resulting system of 4 equations that must be satisfied. Be sure that your first equation is

$$x_1 = (1 - h)F_1x_1 + (1 - h)F_2x_2 + (1 - h)F_3x_3 + (1 - h)F_4x_4. \quad (19.9)$$

- Equation (19.9) as written depends on the entries of the vector \mathbf{x} , but we should be able to arrive at a result that is independent of \mathbf{x} . To see how we do this, we assume the population of the youngest group is never 0, so we can divide both sides of (19.9) by x_1 to obtain

$$1 = (1 - h)F_1 + (1 - h)F_2 \frac{x_2}{x_1} + (1 - h)F_3 \frac{x_3}{x_1} + (1 - h)F_4 \frac{x_4}{x_1}. \quad (19.10)$$

Now we need to write the fractions $\frac{x_2}{x_1}$, $\frac{x_3}{x_1}$, and $\frac{x_4}{x_1}$ so that they do not involve the x_i . Use the remaining equations in your system to show that

$$\begin{aligned} \frac{x_2}{x_1} &= s_1 \\ \frac{x_3}{x_1} &= s_1s_2 \\ \frac{x_4}{x_1} &= s_1s_2s_3. \end{aligned}$$

- Now conclude that the harvesting value h must satisfy the equation

$$1 = (1 - h)[F_1 + F_2s_1 + F_3s_1s_2 + F_4s_1s_2s_3]. \quad (19.11)$$

The value $R = F_1 + F_2s_1 + F_3s_1s_2 + F_4s_1s_2s_3$ is called the *net reproduction rate of the population* and turns out to be the average number of daughters born to a female in her expected lifetime.

- Extend (19.11) to the 12 age group case of the sheep herd. Calculate the value of R for this sheep herd and then find the value of h . Compare this h to the value you obtained through experimentation earlier. Find the fraction of the lambs that should be harvested each year and explain what the stable population state vector \mathbf{x} tells us about the sheep population for this harvesting policy.

Section 20

Complex Eigenvalues

Focus Questions

By the end of this section, you should be able to give precise and thorough answers to the questions listed below. You may want to keep these questions in mind to focus your thoughts as you complete the section.

- What properties do complex eigenvalues of a real matrix satisfy?
- What properties do complex eigenvectors of a real matrix satisfy?
- What is a rotation-scaling matrix?
- How do we find a rotation-scaling matrix within a matrix with complex eigenvalues?

Application: The Gershgorin Disk Theorem

We have now seen different methods for calculating/approximating eigenvalues of a matrix. The algebraic method using the characteristic polynomial can provide exact values, but only in cases where the size of the matrix is small. Methods like the power method allow us to approximate eigenvalues in many, but not all, cases. These approximation techniques can be made more efficient if we have some idea of where the eigenvalues are. The Gershgorin Disc Theorem is a useful tool that can quickly provide bounds on the location of eigenvalues using elementary calculations. For example, using the Gershgorin Disk Theorem we can quickly tell that the real parts of the eigenvalues of the matrix

$$\begin{bmatrix} 3 & 1 & -1 \\ 0 & -1+i & i \\ 2 & 1 & -2i \end{bmatrix}$$

lie between -4 and 5 and the imaginary parts lie between -5 and 2 . Even more, we can say that the eigenvalues lie in the disks (called *Gershgorin disks*) shown in Figure 20.1. We will learn more details about the Gershgorin Disk Theorem at the end of this section.

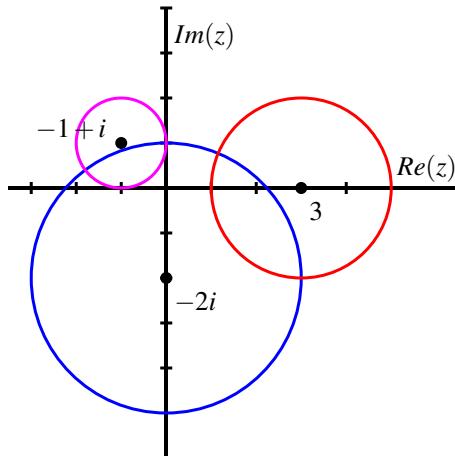


Figure 20.1: Gershgorin disks.

Introduction

So far we have worked with real matrices whose eigenvalues are all real. However, the characteristic polynomial of a matrix with real entries can have complex roots. In this section we investigate the properties of these complex roots and their corresponding eigenvectors, how these complex eigenvectors are found, and the geometric interpretation of the transformations defined by matrices with complex eigenvalues. Although we can consider matrices that have complex numbers as entries, we will restrict ourselves to matrices with real entries.

Preview Activity 20.1. Let $A = \begin{bmatrix} 2 & 4 \\ -2 & 2 \end{bmatrix}$.

- (1) Find the characteristic polynomial of A .
- (2) Find the eigenvalues of A . You should get two complex numbers. How are these complex numbers related?
- (3) Find an eigenvector corresponding to each eigenvalue of A . You should obtain vectors with complex entries.

Complex Eigenvalues

As you noticed in Preview Activity 20.1, the complex roots of the characteristic equation of a real matrix A come in complex conjugate pairs. This should come as no surprise since we know through our use of the quadratic formula that complex roots of (real) quadratic polynomials come in complex conjugate pairs. More generally, if $p(x) = a_0 + a_1x + a_2x^2 + \cdots + a_nx^n$ is a polynomial with real coefficients and z is a root of this polynomial, meaning $p(z) = 0$, then

$$0 = \overline{p(z)} = \overline{a_0 + a_1z + a_2z^2 + \cdots + a_nz^n} = a_0 + a_1\bar{z} + a_2\bar{z}^2 + \cdots + a_n\bar{z}^n = p(\bar{z}).$$

Therefore, \bar{z} is also a root of $p(x)$.

Activity 20.1. Let $A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$.

- (a) The linear transformation $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ defined by $T(\mathbf{x}) = A\mathbf{x}$ is a rotation transformation. What is the angle of rotation?
- (b) Find the eigenvalues of A . For each eigenvalue, find an eigenvector.

In Preview Activity 20.1 and in Activity 20.1, you found that if \mathbf{v} is an eigenvector of A corresponding to λ , then $\bar{\mathbf{v}}$ obtained by taking the complex conjugate of each entry in \mathbf{v} is an eigenvector of A corresponding to $\bar{\lambda}$. Specifically, if $\mathbf{v} = \mathbf{u} + i\mathbf{w}$ where both \mathbf{u} and \mathbf{w} are real vectors is an eigenvector of A , then so is $\bar{\mathbf{v}} = \mathbf{u} - i\mathbf{w}$. We can justify this property using matrix algebra as follows:

$$A\bar{\mathbf{v}} = \overline{A\mathbf{v}} = \overline{A\mathbf{v}} = \overline{\lambda\mathbf{v}} = \bar{\lambda}\bar{\mathbf{v}}.$$

In the first equality, we used the fact that A is a real matrix, so $\overline{A} = A$. In all the other equalities, we used the properties of the conjugation operation in complex numbers.

Rotation and Scaling Matrices

Recall that a rotation matrix is of the form

$$R_\theta = \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix}$$

where the rotation is counterclockwise about the origin by an angle of θ radians. In Activity 20.1, we considered the rotation matrix with angle $\pi/2$ in counterclockwise direction. We will soon see that rotation matrices play an important role in the geometry of a matrix transformation for a matrix that has complex eigenvalues. In this activity, we will restrict ourselves to the 2×2 case, but similar arguments can be made in higher dimensions.

Activity 20.2. Let $A = \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}$.

- (a) Explain why A is not a rotation matrix.
- (b) Although A is not a rotation matrix, there is a rotation matrix B inside A . To find the matrix B , factor out $\sqrt{2}$ from all entries of A . In other words, write A as a product of two matrices in the form

$$A = \begin{bmatrix} \sqrt{2} & 0 \\ 0 & \sqrt{2} \end{bmatrix} B.$$

- (c) The B matrix is a rotation matrix with an appropriate θ . Find this θ .
- (d) If we think about the product of two matrices as applying one transformation after another, describe the effect of the matrix transformation defined by A geometrically.



More generally, if we have a matrix A of the form $A = \begin{bmatrix} a & -b \\ b & a \end{bmatrix}$, then

$$A = \begin{bmatrix} \sqrt{a^2 + b^2} & 0 \\ 0 & \sqrt{a^2 + b^2} \end{bmatrix} \begin{bmatrix} \frac{a}{\sqrt{a^2 + b^2}} & \frac{-b}{\sqrt{a^2 + b^2}} \\ \frac{b}{\sqrt{a^2 + b^2}} & \frac{a}{\sqrt{a^2 + b^2}} \end{bmatrix}.$$

The first matrix in the decomposition is a scaling matrix with a scaling factor of $s = \sqrt{a^2 + b^2}$. So if $s > 1$, the transformation stretches vectors, and if $s < 1$, the transformation shrinks vectors. The second matrix in the decomposition is a rotation matrix with angle θ such that $\cos(\theta) = \frac{a}{\sqrt{a^2 + b^2}}$ and $\sin(\theta) = \frac{b}{\sqrt{a^2 + b^2}}$. This angle is also the angle between the positive x -axis and the vector $\mathbf{v} = \begin{bmatrix} a \\ b \end{bmatrix}$. We will refer to the matrices of the form $\begin{bmatrix} a & -b \\ b & a \end{bmatrix}$ as *rotation-scaling matrices*.

Matrices with Complex Eigenvalues

Now we will investigate how a general 2×2 matrix with complex eigenvalues can be seen to be similar (both in a linear algebra and a colloquial meaning) to a rotation-scaling matrix.

Activity 20.3. Let $B = \begin{bmatrix} 1 & -5 \\ 2 & 3 \end{bmatrix}$. The eigenvalues of B are $2 \pm 3i$. An eigenvector for the eigenvalue $2 - 3i$ is $\mathbf{v} = \begin{bmatrix} -5 \\ 1 - 3i \end{bmatrix}$. We will use this eigenvector to show that B is similar to a rotation-scaling matrix.

- (a) Any complex vector \mathbf{v} can be written as $\mathbf{v} = \mathbf{u} + i\mathbf{w}$ where both \mathbf{u} and \mathbf{w} are real vectors. What are these real vectors \mathbf{u} and \mathbf{w} for the eigenvector \mathbf{v} above?
- (b) Let $P = [\mathbf{u} \ \mathbf{w}]$ be the matrix whose first column is the real part of \mathbf{v} and whose second column is the imaginary part of \mathbf{v} (without the i). Find $R = P^{-1}BP$.
- (c) Express R as a product of a rotation and a scaling matrix. What is the factor of scaling? What is the rotation angle?

In Activity 20.3, we saw that the matrix B with complex eigenvalues $2 \pm 3i$ is similar to a rotation-scaling matrix. Specifically $R = P^{-1}BP$, where the columns of P are the real and imaginary parts of an eigenvector of B , is the rotation-scaling matrix with a factor of scaling by $\sqrt{2^2 + 3^2}$ and a rotation by angle $\theta = \arccos\left(\frac{2}{\sqrt{2^2 + 3^2}}\right)$.

Does a similar decomposition result hold for a general 2×2 matrix with complex eigenvalues? We investigate this question in the next activity.

Activity 20.4. Let A be a 2×2 matrix with complex eigenvalue $\lambda = a - bi$, $b \neq 0$, and corresponding complex eigenvector $\mathbf{v} = \mathbf{u} + i\mathbf{w}$.

- (a) Explain why $A\mathbf{v} = A\mathbf{u} + iA\mathbf{w}$.
- (b) Explain why $\lambda\mathbf{v} = (a\mathbf{u} + b\mathbf{w}) + i(a\mathbf{w} - b\mathbf{u})$.



(c) Use the previous two results to explain why

- $A\mathbf{u} = a\mathbf{u} + b\mathbf{w}$ and
- $A\mathbf{w} = a\mathbf{w} - b\mathbf{u}$.

(d) Let $P = [\mathbf{u} \ \mathbf{w}]$. We will now show that $AP = PR$ where $R = \begin{bmatrix} a & -b \\ b & a \end{bmatrix}$.

i. Without any calculation, explain why

$$AP = [A\mathbf{u} \ A\mathbf{w}].$$

ii. Recall that if M is an $m \times n$ matrix and \mathbf{x} is an $n \times 1$ vector, then the matrix product $M\mathbf{x}$ is a linear combination of the columns of M with weights the corresponding entries of the vector \mathbf{x} . Use this idea to show that

$$PR = [a\mathbf{u} + b\mathbf{w} \ -b\mathbf{u} + a\mathbf{w}].$$

iii. Now explain why $AP = PR$.

iv. Assume for the moment that P is an invertible matrix. Show that $A = PRP^{-1}$.

Your work in Activity 20.4 shows that any 2×2 matrix is similar to a rotation-scaling matrix with a factor of scaling by $\sqrt{a^2 + b^2}$ and a rotation by angle $\theta = \arccos(\frac{a}{\sqrt{a^2 + b^2}})$ if $b \geq 0$, and $\theta = -\arccos(\frac{a}{\sqrt{a^2 + b^2}})$ if $b < 0$. Geometrically, this means that every 2×2 real matrix with complex eigenvalues is just a scaled rotation (R) with respect to the basis \mathcal{B} formed by \mathbf{u} and \mathbf{w} from the complex eigenvector \mathbf{v} . Multiplying by P^{-1} and P simply provides the change of basis from the standard basis to the basis \mathcal{B} , as we will see in detail when we learn about linear transformations.

Theorem 20.1. *Let A be a real 2×2 matrix with complex eigenvalue $a - bi$ and corresponding eigenvector $\mathbf{v} = \mathbf{u} + i\mathbf{w}$. Then*

$$A = PRP^{-1}, \text{ where } P = [\mathbf{u} \ \mathbf{w}] \text{ and } R = \begin{bmatrix} a & -b \\ b & a \end{bmatrix}.$$

The one fact that we have not yet addressed is why the matrix $P = [\mathbf{u} \ \mathbf{w}]$ is invertible. We do that now to complete the argument.

Let A be a real 2×2 matrix with $A\mathbf{v} = \lambda\mathbf{v}$, where $\lambda = a - bi$, $b \neq 0$ and $\mathbf{v} = \mathbf{u} + i\mathbf{w}$. To show that \mathbf{u} and \mathbf{w} are linearly independent, we need to show that no nontrivial linear combination of \mathbf{u} and \mathbf{w} can be the zero vector. Suppose

$$x_1\mathbf{u} + x_2\mathbf{w} = \mathbf{0}$$

for some scalars x_1 and x_2 . We will show that $x_1 = x_2 = 0$. Assume to the contrary that one of x_1, x_2 is not zero. First, assume $x_1 \neq 0$. Then $\mathbf{u} = -\frac{x_2}{x_1}\mathbf{w}$. Let $c = -\frac{x_2}{x_1}$. Then

$$\begin{aligned} A\mathbf{u} &= A(c\mathbf{w}) \\ A\mathbf{u} &= cA\mathbf{w} \\ a\mathbf{u} + b\mathbf{w} &= c(a\mathbf{u} - b\mathbf{w}) \\ (a + cb)\mathbf{u} &= (ca - b)\mathbf{w} \\ (a + cb)(c\mathbf{u}) &= (ca - b)\mathbf{w}. \end{aligned}$$



So we must have $(a+cb)c = ca - b$. This equation simplifies to $c^2b = -b$. Since $b \neq 0$, we conclude that $c^2 = -1$ which is impossible for a real constant c . Therefore, we cannot have $x_1 \neq 0$. A similar argument (left to the reader) shows that $x_2 = 0$. Thus we can conclude that \mathbf{u} and \mathbf{w} are linearly independent.

Examples

What follows are worked examples that use the concepts from this section.

Example 20.2. Let $A = \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & -1 \\ 1 & 1 & 1 \end{bmatrix}$.

- Without doing any computations, explain why not all of the eigenvalues of A can be complex.
- Find all of the eigenvalues of A .

Example Solution.

- Since complex eigenvalues occur in conjugate pairs, the complex eigenvalues with nonzero imaginary parts occur in pairs. Since A can have at most 3 different eigenvalues, at most two of them can have nonzero imaginary parts. So at least one eigenvalue of A is real.

(b) For this matrix A we have $A - \lambda I_3 = \begin{bmatrix} -\lambda & 1 & 0 \\ -1 & -\lambda & -1 \\ 1 & 1 & -\lambda + 1 \end{bmatrix}$. Using a cofactor expansion along the first row gives us

$$\begin{aligned} \det(A - \lambda I_3) &= (-\lambda)((-\lambda)(1 - \lambda) + 1) - ((-1)(1 - \lambda) + 1) \\ &= -\lambda^3 + \lambda^2 - \lambda + 1 - \lambda - 1 \\ &= \lambda^3 + \lambda^2 - 2\lambda \\ &= -\lambda(\lambda^2 - \lambda + 2). \end{aligned}$$

The roots of the characteristic polynomial are $\lambda = 0$ and

$$\lambda = \frac{1 \pm \sqrt{1 - 4(2)}}{2} = \frac{1}{2}(1 \pm \sqrt{7}i).$$

Example 20.3. Let $A = \begin{bmatrix} 1 & 2 \\ -1 & 3 \end{bmatrix}$. Find a rotation scaling matrix R that is similar to A . Identify the rotation and scaling factor.

Example Solution. The eigenvalues of A are the roots of the characteristic polynomial

$$\begin{aligned} p(\lambda) &= \det(A - \lambda I_2) \\ &= \det \left(\begin{bmatrix} 1 - \lambda & 2 \\ -1 & 3 - \lambda \end{bmatrix} \right) \\ &= (1 - \lambda)(3 - \lambda) + 2 \\ &= \lambda^2 - 4\lambda + 5. \end{aligned}$$

The quadratic formula shows that the roots of $p(\lambda)$ are

$$\frac{4 \pm \sqrt{-4}}{2} = 2 \pm i.$$

To find an eigenvector for A with eigenvalue $2 - i$, we row reduce

$$A - (2 - i)I_3 = \begin{bmatrix} -1 + i & 2 \\ -1 & 1 + i \end{bmatrix}$$

to

$$\begin{bmatrix} 1 & -i - 1 \\ 0 & 0 \end{bmatrix}.$$

An eigenvector for A with eigenvalue $2 - i$ is then

$$[1+i \ 1]^T = [1 \ 1]^T + i[1 \ 0]^T.$$

Letting $P = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}$, we have

$$R = P^{-1}AP = \begin{bmatrix} 2 & -1 \\ 1 & 2 \end{bmatrix}.$$

The scaling is determined by the determinant of R which is 5, and the angle θ of rotation satisfies $\sin(\theta) = \frac{1}{5}$. This makes $\theta \approx 0.2014$ radians or approximately 11.5370° counterclockwise.

Summary

- For a real matrix, complex eigenvalues appear in conjugate pairs. Specifically, if $\lambda = a + ib$ is an eigenvalue of a real matrix A , then $\bar{\lambda} = a - ib$ is also an eigenvalue of A .
- For a real matrix, if a \mathbf{v} is an eigenvector corresponding to λ , then the vector $\bar{\mathbf{v}}$ obtained by taking the complex conjugate of each entry in \mathbf{v} is an eigenvector corresponding to $\bar{\lambda}$.
- The rotation-scaling matrix $A = \begin{bmatrix} a & -b \\ b & a \end{bmatrix}$ can be written as

$$\begin{bmatrix} \sqrt{a^2 + b^2} & 0 \\ 0 & \sqrt{a^2 + b^2} \end{bmatrix} \begin{bmatrix} \frac{a}{\sqrt{a^2 + b^2}} & \frac{-b}{\sqrt{a^2 + b^2}} \\ \frac{b}{\sqrt{a^2 + b^2}} & \frac{a}{\sqrt{a^2 + b^2}} \end{bmatrix}.$$

This decomposition geometrically means that the transformation corresponding to A can be viewed as a rotation by angle $\theta = \arccos\left(\frac{a}{\sqrt{a^2+b^2}}\right)$ if $b \geq 0$, or $\theta = -\arccos\left(\frac{a}{\sqrt{a^2+b^2}}\right)$ if $b < 0$, followed by a scaling by factor $\sqrt{a^2 + b^2}$.

- If A is a real 2×2 matrix with complex eigenvalue $a - bi$ and corresponding eigenvector $\mathbf{v} = \mathbf{u} + i\mathbf{w}$, then A is similar to the rotation-scaling matrix $R = \begin{bmatrix} a & -b \\ b & a \end{bmatrix}$. More specifically,

$$A = P R P^{-1}, \text{ where } P = [\mathbf{u} \ \mathbf{w}].$$



Exercises

(1) Find eigenvalues and eigenvectors of each of the following matrices.

(a) $\begin{bmatrix} 2 & 4 \\ -2 & 2 \end{bmatrix}$

(b) $\begin{bmatrix} 3 & 2 \\ -1 & 1 \end{bmatrix}$

(c) $\begin{bmatrix} 1 & -2 \\ 4 & -3 \end{bmatrix}$

(2) Find a rotation-scaling matrix where the rotation angle is $\theta = 3\pi/4$ and scaling factor is less than 1.

(3) Determine which rotation-scaling matrices have determinant equal to 1. Be as specific as possible.

(4) Determine the rotation-scaling matrix inside the matrix $\begin{bmatrix} 2 & 4 \\ -2 & 2 \end{bmatrix}$.

(5) Find a real 2×2 matrix with eigenvalue $1 + 2i$.

(6) Find a real 2×2 matrix which is not a rotation-scaling matrix with eigenvalue $-1 + 2i$.

(7) We have seen how to find the characteristic polynomial of an $n \times n$ matrix. In this exercise we consider the reverse question. That is, given a polynomial $p(\lambda)$ of degree n , can we find an $n \times n$ matrix whose characteristic polynomial is $p(\lambda)$?

(a) Find the characteristic polynomial of the 2×2 matrix $C = \begin{bmatrix} 0 & -a_0 \\ 1 & -a_1 \end{bmatrix}$. Use this result to find a real valued matrix whose eigenvalues are $1 + i$ and $1 - i$.

(b) Repeat part (a) by showing that $-p(\lambda) = -(\lambda^2 + a_2\lambda^2 + a_1\lambda + a_0)$ is the characteristic polynomial of the 3×3 matrix $C = \begin{bmatrix} 0 & 0 & -a_0 \\ 0 & 1 & -a_1 \\ 0 & 0 & -a_2 \end{bmatrix}$.

(c) We can generalize this argument. Prove, using mathematical induction, that the polynomial

$$p(\lambda) = (-1)^n (\lambda^n + a_{n-1}\lambda^{n-1} + a_{n-2}\lambda^{n-2} + \cdots + a_1\lambda + a_0)$$

is the characteristic polynomial of the matrix

$$C = \begin{bmatrix} 0 & 0 & 0 & \cdots & 0 & -a_0 \\ 1 & 0 & 0 & \cdots & 0 & -a_1 \\ 0 & 1 & 0 & \cdots & 0 & -a_2 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & -a_{n-1} \end{bmatrix}.$$

The matrix C is called the *companion matrix* for $p(\lambda)$.



- (8) Label each of the following statements as True or False. Provide justification for your response.
- True/False** If $3 - 4i$ is an eigenvalue of a real matrix, then so is $3 + 4i$.
 - True/False** If $2 + 3i$ is an eigenvalue of a 3×3 real matrix A , then A has three distinct eigenvalues.
 - True/False** Every 2×2 real matrix with complex eigenvalues is a rotation-scaling matrix.
 - True/False** Every square matrix with real entries has real number eigenvalues.
 - True/False** If A is a 2×2 matrix with complex eigenvalues similar to a rotation-scaling matrix R , the eigenvalues of A and R are the same.
 - True/False** If A is a real matrix with complex eigenvalues, all eigenvectors of A must be non-real.

Project: Understanding the Gershgorin Disk Theorem

To understand the Gershgorin Disk Theorem, we need to recall how to visualize a complex number in the plane. Recall that a complex number z is a number of the form $z = a + bi$ where a and b are real numbers and $i^2 = -1$. The number a is the real part of z , denoted as $\Re(z)$, and b is the imaginary part of z , denoted $\Im(z)$. The set of all complex numbers is denoted \mathbb{C} . We define addition and multiplication on \mathbb{C} as follows. For $a + bi, c + di \in \mathbb{C}$,

$$(a + bi) + (c + di) = (a + c) + (b + d)i \quad \text{and} \quad (a + bi)(c + di) = (ac - bd) + (ad + bc)i.$$

Note that the product is what we would expect if we “expanded” the product in the normal way and used the fact that $i^2 = -1$. The set of complex numbers forms a field – that is, \mathbb{C} satisfies all of the same properties as \mathbb{R} as stated in Theorem 4.2.

We can visualize the complex number $a + bi$ in the plane as the point (a, b) . Here we are viewing the horizontal axis as the real axis and the vertical axis as the imaginary axis. The length (or magnitude) of the complex number $z = a + bi$, which we denote as $|z|$, is the distance from the origin to z . So by the Pythagorean Theorem we have $|a + bi| = \sqrt{a^2 + b^2}$. Note that the magnitude of $z = a + bi$ can be written as a complex product

$$|z| = \sqrt{(a + bi)(a - bi)}.$$

The complex number $a - bi$ is called the *complex conjugate* of $z = a + bi$ and is denoted as \bar{z} . A few important properties of real numbers and their conjugates are the following. Let $z = a + bi$ and $w = c + di$ be complex numbers. Then

- $\overline{z + w} = \overline{(a + c) + (b + d)i} = (a + c) - (b + d)i = (a - bi) + (c - di) = \bar{z} + \bar{w}$,
- $\overline{zw} = \overline{(ac - bd) + (ad + bc)i} = (ac - bd) - (ad + bc)i = (a - bi)(c - di) = \bar{z}\bar{w}$,
- $\overline{\bar{z}} = z$,



- $|z| = \sqrt{a^2 + b^2} \geq \sqrt{a^2} = |a| = |\Re(z)|$,
- $|z| = \sqrt{a^2 + b^2} \geq \sqrt{b^2} = |b| = |\Im(z)|$,
- $|\bar{z}| = |z|$,
- $|z| = 0$ if and only if $z = 0$,
- If $p(x)$ is a polynomial with real coefficients and the complex number z satisfies $p(z) = 0$, then $p(\bar{z}) = 0$ as well.

Using these facts we can show that the triangle inequality is true for complex numbers. That is,

$$|z + w| \leq |z| + |w|.$$

To see why, notice that

$$\begin{aligned} |z + w|^2 &= (z + w)(\overline{z + w}) \\ &= (z + w)(\bar{z} + \bar{w}) \\ &= z\bar{z} + z\bar{w} + w\bar{z} + w\bar{w} \\ &= z\bar{z} + z\bar{w} + \overline{zw\bar{w}} + w\bar{w} \\ &= |z|^2 + 2\Re(z\bar{w}) + |w|^2 \\ &\leq |z|^2 + 2|zw| + |w|^2 \\ &= |z|^2 + 2|z||w| + |w|^2 \\ &= (|z| + |w|)^2. \end{aligned}$$

Since $|z + w|$, $|z|$, and $|w|$ are all non-negative, taking square roots of both sides gives us $|z + w| \leq |z| + |w|$ as desired. We can extend this triangle inequality to any number of complex numbers. That is, if z_1, z_2, \dots, z_k are complex numbers, then

$$|z_1 + z_2 + \dots + z_k| \leq |z_1| + |z_2| + \dots + |z_k|. \quad (20.1)$$

We can prove Equation (20.1) by mathematical induction. We have already done the $k = 2$ case and so we assume that Equation (20.1) is true for any sum of k complex numbers. Now suppose that $z_1, z_2, \dots, z_k, z_{k+1}$ are complex numbers. Then

$$\begin{aligned} |z_1 + z_2 + \dots + z_k + z_{k+1}| &= |(z_1 + z_2 + \dots + z_k) + z_{k+1}| \\ &\leq |z_1 + z_2 + \dots + z_k| + |z_{k+1}| \\ &\leq (|z_1| + |z_2| + \dots + |z_k|) + |z_{k+1}| \\ &= |z_1| + |z_2| + \dots + |z_k| + |z_{k+1}|. \end{aligned}$$

To prove the Gershgorin Disk Theorem, we will use the Levy-Desplanques Theorem, which gives conditions that guarantee that a matrix is invertible. We illustrate with an example in the following activity.

Project Activity 20.1. Let $A = \begin{bmatrix} 3 & 2 \\ -1 & 4 \end{bmatrix}$. Since $\det(A) \neq 0$, we know that A is an invertible matrix. Let us assume for a moment that we don't know that A is invertible and try to determine



if 0 is an eigenvalue of A . In other words, we want to know if there is a nonzero vector \mathbf{v} so that $A\mathbf{v} = \mathbf{0}$. Assuming the existence of such a vector $\mathbf{v} = [v_1 \ v_2]^\top$, for $A\mathbf{v}$ to be $\mathbf{0}$ it must be the case that

$$3v_1 + 2v_2 = 0 \text{ and } -v_1 + 4v_2 = 0.$$

Since the vector \mathbf{v} is not the zero vector, at least one of v_1, v_2 is not zero. Note that if one of v_1, v_2 is zero, the so is the other. So we can assume that v_1 and v_2 are nonzero.

- (a) Use the fact that $3v_1 + 2v_2 = 0$ to show that $|v_2| > |v_1|$.
- (b) Use the fact that $-v_1 + 4v_2 = 0$ to show that $|v_1| > |v_2|$. What conclusion can we draw about whether 0 is an eigenvalue of A ? Why does this mean that A is invertible?

What makes the arguments work in Project Activity 20.1 is that $|3| > |2|$ and $|4| > |-1|$. This argument can be extended to larger matrices, as described in the following theorem.

Theorem 20.4 (Levy-Desplanques Theorem). *Any square matrix $A = [a_{ij}]$ satisfying $|a_{ii}| > \sum_{j \neq i} |a_{ij}|$ for all i is invertible.*

Proof. Let $A = [a_{ij}]$ be an $n \times n$ matrix satisfying $|a_{ii}| > \sum_{j \neq i} |a_{ij}|$ for all i . Let us assume that A is not invertible, that is that there is a vector $\mathbf{v} \neq \mathbf{0}$ such that $A\mathbf{v} = \mathbf{0}$. Let $\mathbf{v} = [v_1 \ v_2 \ \dots \ v_n]$ and t be between 1 and n so that $|v_t| \geq |v_i|$ for all i . That is, choose v_t to be the component of \mathbf{v} with the largest absolute value.

Expanding the product $A\mathbf{v}$ using the row-column product along the t th row shows that

$$a_{t1}v_1 + a_{t2}v_2 + \dots + a_{tn}v_n = 0.$$

Solving for the a_{tt} term gives us

$$a_{tt}v_t = -(a_{t1}v_1 + a_{t2}v_2 + \dots + a_{t(t-1)}v_{t-1} + a_{t(t+1)}v_{t+1} + \dots + a_{tn}v_n).$$

Then

$$\begin{aligned} |a_{tt}| |v_t| &= |-(a_{t1}v_1 + a_{t2}v_2 + \dots + a_{t(t-1)}v_{t-1} + a_{t(t+1)}v_{t+1} + \dots + a_{tn}v_n)| \\ &= |a_{t1}v_1 + a_{t2}v_2 + \dots + a_{t(t-1)}v_{t-1} + a_{t(t+1)}v_{t+1} + \dots + a_{tn}v_n| \\ &\leq |a_{t1}| |v_1| + |a_{t2}| |v_2| + \dots + |a_{t(t-1)}| |v_{t-1}| + |a_{t(t+1)}| |v_{t+1}| + \dots + |a_{tn}| |v_n| \\ &\leq |a_{t1}| |v_t| + |a_{t2}| |v_t| + \dots + |a_{t(t-1)}| |v_t| + |a_{t(t+1)}| |v_t| + \dots + |a_{tn}| |v_t| \\ &= (|a_{t1}| + |a_{t2}| + \dots + |a_{t(t-1)}| + |a_{t(t+1)}| + \dots + |a_{tn}|) |v_t|. \end{aligned}$$

Since $|v_t| \neq 0$, we cancel the $|v_t|$ term to conclude that

$$|a_{tt}| \leq |a_{t1}| + |a_{t2}| + \dots + |a_{t(t-1)}| + |a_{t(t+1)}| + \dots + |a_{tn}|.$$

But this contradicts the condition that $|a_{ii}| > \sum_{j \neq i} |a_{ij}|$ for all i . We conclude that 0 is not an eigenvalue for A and A is invertible. \blacksquare

Any matrix $A = [a_{ij}]$ satisfying the condition of the Levy-Desplanques Theorem is given a special name.



Definition 20.5. A square matrix $A = [a_{ij}]$ is **strictly diagonally dominant** if $|a_{ii}| > \sum_{j \neq i} |a_{ij}|$ for all i .

So any strictly diagonally dominant matrix is invertible. A quick glance can show that a matrix is strictly diagonally dominant. For example, since $|3| > |1| + |-1|$, $|12| > |5| + |6|$, and $|-8| < |-2| + |4|$, the matrix

$$A = \begin{bmatrix} 3 & 1 & -1 \\ 5 & 12 & 6 \\ -2 & 4 & -8 \end{bmatrix}$$

is strictly diagonally dominant and therefore invertible. However, just because a matrix is not strictly diagonally dominant, it does not follow that the matrix is non-invertible. For example, the matrix $B = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}$ is invertible, but not strictly diagonally dominant.

Now we can address the Gershgorin Disk Theorem.

Project Activity 20.2. Let A be an arbitrary $n \times n$ matrix and assume that λ is an eigenvalue of A .

- (a) Explain why the matrix $A - \lambda I$ is singular.
- (b) What does the Levy-Desplanques Theorem tell us about the matrix $A - \lambda I$?
- (c) Explain how we can conclude the Gershgorin Disk Theorem.

Theorem 20.6 (Gershgorin Disk Theorem). *Let $A = [a_{ij}]$ be an $n \times n$ matrix with complex entries. Then every eigenvalue of A lies in one of the Gershgorin discs*

$$\{z \in \mathbb{C} : |z - a_{ii}| \leq r_i\},$$

where $r_i = \sum_{j \neq i} |a_{ij}|$.

Based on this theorem, we define a Gershgorin disk to be $D(a_{ii}, r_i)$, where $r_i = \sum_{j \neq i} |a_{ij}|$.

- (d) Use the Gershgorin Disk Theorem to give estimates on the locations of the eigenvalues of the matrix $A = \begin{bmatrix} -1 & 2 \\ -3 & 2 \end{bmatrix}$.

The Gershgorin Disk Theorem has a consequence that gives additional information about the eigenvalues if some of the Gershgorin disks do not overlap.

Theorem 20.7. *If S is a union of m Gershgorin disks of a matrix A such that S does not intersect any other Gershgorin disk, then S contains exactly m eigenvalues (counting multiplicities) of A .*

Proof. Most proofs of this theorem require some results from topology. For that reason, we will not present a completely rigorous proof but rather give the highlights. Let $A = [a_{ij}]$ be an $n \times n$ matrix. Let D_i be a collection of Gershgorin disks of A for $1 \leq i \leq m$ such that $S = \cup_{1 \leq i \leq m} D_i$ does not intersect any other Gershgorin disk of A , and let S' be the union of the Gershgorin disks of A that are different from the D_i . Note that $S \cap S' = \emptyset$. Let C be the matrix whose i th column is $a_{ii}\mathbf{e}_i$, that is C is the diagonal matrix whose diagonal entries are the corresponding diagonal entries



of A . Note that the eigenvalues of C are a_{ii} and the Gershgorin disks of C are just the points a_{ii} . So our theorem is true for this matrix C . To prove the result, we build a continuum of matrices from C to A as follows: let $B = A - C$ (so that B is the matrix whose off-diagonal entries are those of A and whose diagonal entries are 0), and let $A(t) = tB + C$ for t in the interval $[0, 1]$. Note that $A(1) = A$. Since the diagonal entries of $A(t)$ are the same as those of A , the Gershgorin disks of $A(t)$ have the same centers as the corresponding Gershgorin disks of A , while the radii of the Gershgorin disks of $A(t)$ are those of A but scaled by t . So the Gershgorin disks of $A(t)$ increase from points (the a_{ii}) to the Gershgorin disks of A as t increases from 0 to 1. While the centers of the disks all remain fixed, it is important to recognize that the eigenvalues of $A(t)$ move as t changes. An illustration of this is shown in Figure 20.2 with the eigenvalues as the black points and the changing Gershgorin disks dashed in magenta, using the matrix $\begin{bmatrix} i & \frac{1}{2} \\ 1 & -2+i \end{bmatrix}$. We can learn about how the eigenvalues move with the characteristic polynomial.

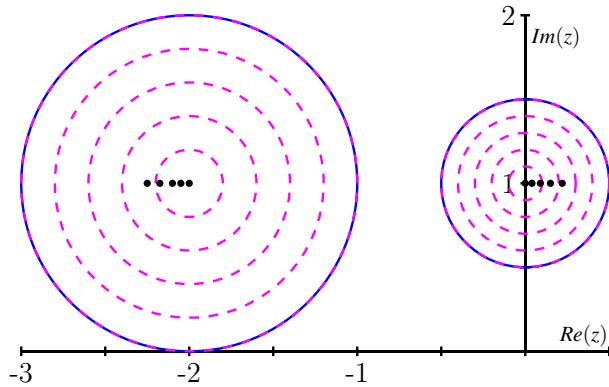


Figure 20.2: How eigenvalues move.

Let $p(t, x)$ be the characteristic polynomial of $A(t)$. Note that these characteristic polynomials are functions of both t and x . Since polynomials are continuous functions, their roots (the eigenvalues of $A(t)$) are continuous for $t \in [0, 1]$ as well. Let $\lambda(t)$ be an eigenvalue of $A(t)$. Note that $\lambda(1)$ is an eigenvalue of A , and $\lambda(0)$ is one of the a_{ii} and is therefore in S . We will argue that $\lambda(t)$ is in S for every value of t in $[0, 1]$. Let r_i be the radius of D_i and let $D(t)_i$ be the Gershgorin disk of $A(t)$ with the same center as D_i and radius $r(t)_i = tr_i$. Let $S(t) = \cup_{1 \leq i \leq m} D(s)_i$. Since $r(s)_i \leq r_i$, it follows that $D(s)_i \subseteq D_i$ and so $S(t) \cap S' = \emptyset$ as well. From topology, we know that since the disks D_i are closed, the union S of these disks is also closed. Similarly, $S(t)$ and S' are closed. Thus, $\lambda(t)$ is continuous in a closed set and so does not leave the set. Thus, $\lambda(t)$ is in S for every value of t in $[0, 1]$. ■

Section 21

Properties of Determinants

Focus Questions

By the end of this section, you should be able to give precise and thorough answers to the questions listed below. You may want to keep these questions in mind to focus your thoughts as you complete the section.

- How do elementary row operations change the determinant?
- How can we represent elementary row operations via matrix multiplication?
- How can we use elementary row operations to calculate the determinant more efficiently?
- What is the Cramer's rule for the explicit formula for the inverse of a matrix?
- How can we interpret determinants from a geometric perspective?

Introduction

This section is different than others in that it contains mainly proofs of previously stated results and only a little new material. Consequently, there is no application attached to this section.

We have seen that an important property of the determinant is that it provides an easy criteria for the invertibility of a matrix. As a result, we obtained an algebraic method for finding the eigenvalues of a matrix, using the characteristic equation. In this section, we will investigate other properties of the determinant related to how elementary row operations change the determinant. These properties of the determinant will help us evaluate the determinant in a more efficient way compared to using the cofactor expansion method, which is computationally intensive for large n values due to it being a recursive method. Finally, we will derive a geometrical interpretation of the determinant.

Preview Activity 21.1.

- (1) We first consider how the determinant changes if we multiply a row of the matrix by a constant.

(a) Let $A = \begin{bmatrix} 2 & 3 \\ 1 & 4 \end{bmatrix}$. Pick a few different values for the constant k and compare the determinant of A and that of $\begin{bmatrix} 2k & 3k \\ 1 & 4 \end{bmatrix}$. What do you conjecture that the effect of multiplying a row by a constant on the determinant is?

- (b) If we want to make sure our conjecture is valid for any 2×2 matrix, we need to show that for $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, the relationship between $\det(A)$ and the determinant of $\begin{bmatrix} a \cdot k & b \cdot k \\ c & d \end{bmatrix}$ follows our conjecture. We should also check that the relationship between $\det(A)$ and the determinant of $\begin{bmatrix} a & b \\ c \cdot k & d \cdot k \end{bmatrix}$ follows our conjecture. Verify this.
- (c) Make a similar conjecture for what happens to the determinant when a row of a 3×3 matrix A is multiplied by a constant k , and explain why your conjecture is true using the cofactor expansion definition of the determinant.

- (2) The second type of elementary row operation we consider is row swapping.

(a) Take a general 2×2 matrix $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ and determine how row swapping effects the determinant.

(b) Now choose a few different 3×3 matrices and see how row swapping changes the determinant in these matrices by evaluating the determinant with a calculator or any other appropriate technology.

(c) Based on your results so far, conjecture how row swapping changes the determinant in general.

- (3) The last type of elementary row operation is adding a multiple of a row to another. Determine the effect of this operation on a 2×2 matrix by evaluating the determinant of a general 2×2 matrix after a multiple of one row is added to the other row.

- (4) All of the elementary row operations we discussed above can be achieved by matrix multiplication with *elementary matrices*. For each of the following elementary matrices, determine what elementary operation it corresponds to by calculating the product EA , where

$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$ is a general 3×3 matrix.

$$(a) E = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (b) E = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (c) E = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix}$$



Elementary Row Operations and Their Effects on the Determinant

In Preview Activity 21.1, we conjectured how elementary row operations affect the determinant of a matrix. In the following activity, we prove how the determinant changes when a row is multiplied by a constant using the cofactor expansion definition of the determinant.

Activity 21.1. In this activity, assume that the determinant of A can be determined by a cofactor expansion along any row or column. (We will prove this result independently later in this section.) Consider an arbitrary $n \times n$ matrix $A = [a_{ij}]$.

- (a)
- (b) Write the expression for $\det(A)$ using the cofactor expansion along the second row.
- (c) Let B be obtained by multiplying the second row of A by k . Write the expression for $\det(B)$ if the cofactor expansion along the second row is used.
- (d) Use the expressions you found above, to express $\det(B)$ in terms of $\det(A)$.
- (e) Explain how this method generalizes to prove the relationship between the determinant of a matrix A and that of the matrix obtained by multiplying a row by a constant k .

Your work in Activity 21.1 proves the first part of the following theorem on how elementary row operations change the determinant of a matrix.

Theorem 21.1. *Let A be a square matrix.*

- (1) *If B is obtained by multiplying a row of A by a constant k , then $\det(B) = k \det(A)$.*
- (2) *If B is obtained by swapping two rows of A , then $\det(B) = -\det(A)$.*
- (3) *If B is obtained by adding a multiple of a row of A to another, then $\det(B) = \det(A)$.*

In the next section, we will use elementary matrices to prove the last two properties of Theorem 21.1.

Elementary Matrices

As we saw in Preview Activity 21.1, elementary row operations can be achieved by multiplication by *elementary matrices*.

Definition 21.2. An **elementary matrix** is a matrix obtained by performing a single elementary row operation on an identity matrix.

The following elementary matrices correspond, respectively, to an elementary row operation which swaps rows 2 and 4; an elementary row operation which multiplies the third row by 5; and an elementary row operation which adds four times the third row to the first row on any 4×4 matrix:



$$E_1 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}, \quad E_2 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 5 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad \text{and} \quad E_3 = \begin{bmatrix} 1 & 0 & 4 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

To obtain an elementary matrix corresponding an elementary row operation, we simply perform the elementary row operation on the identity matrix. For example, E_1 above is obtained by swapping rows 2 and 4 of the identity matrix.

With the use of elementary matrices, we can now prove the result about how the determinant is affected by elementary row operations. We first rewrite Theorem 21.1 in terms of elementary matrices:

Theorem 21.3. *Let A be an $n \times n$ matrix. If E is an $n \times n$ elementary matrix, then $\det(EA) = \det(E) \det(A)$ where*

$$\det(E) = \begin{cases} r & \text{if } E \text{ corresponds to multiplying a row by } r \\ -1 & \text{if } E \text{ corresponds to swapping two rows} \\ 1 & \text{if } E \text{ corresponds to adding a multiple of a row to another.} \end{cases}$$

Notes on Theorem 21.3. An elementary matrix E obtained by multiplying a row by r is a diagonal matrix with one r along the diagonal and the rest 1s, so $\det(E) = r$. Similarly, an elementary matrix E obtained by adding a multiple of a row to another is a triangular matrix with 1s along the diagonal, so $\det(E) = 1$. The fact that the the determinant of an elementary matrix obtained by swapping two rows is -1 is a bit more complicated and is verified independently later in this section. Also, the proof of 21.3 depends on the fact that the cofactor expansion of a matrix is the same along any two rows. A proof of this can also be found later in this section.

Proof of Theorem 21.3. We will prove the result by induction on n , the size of the matrix A . We verified these results in Preview Activity 21.1 for $n = 2$ using elementary row operations. The elementary matrix versions follow immediately.

Now assume the theorem is true for $k \times k$ matrices with $k \geq 2$ and consider an $n \times n$ matrix A where $n = k + 1$. If E is an $n \times n$ elementary matrix, we want to show that $\det(EA) = \det(E) \det(A)$. Let $EA = B$. (Although it is an abuse of language, we will refer to both the elementary matrix and the elementary row operation corresponding to it by E .)

When finding $\det(B) = \det(EA)$ we will use a cofactor expansion along a row which is not affected by the elementary row operation E . Since E affects at most two rows and A has $n \geq 3$ rows, it is possible to find such a row, say row i . The cofactor expansion along row i of B is

$$b_{i1}(-1)^{i+1} \det(B_{i1}) + b_{i2}(-1)^{i+2} \det(B_{i2}) + \cdots + b_{in}(-1)^{i+n} \det(B_{in}). \quad (21.1)$$

Since we chose a row of A which was not affected by the elementary row operation, it follows that $b_{ij} = a_{ij}$ for $1 \leq j \leq n$. Also, the matrix B_{ij} obtained by removing row i and column j from matrix $B = EA$ can be obtained from A_{ij} by an elementary row operation of the same type as E . Hence there is an elementary matrix E_k of the same type as E with $B_{ij} = E_k A_{ij}$. Therefore, by



induction, $\det(B_{ij}) = \det(E_k) \det(A_{ij})$ and $\det(E_k)$ is equal to 1, -1 or r depending on the type of elementary row operation. If we substitute this information into equation (21.1), we obtain

$$\begin{aligned}\det(B) &= a_{i1}(-1)^{i+1} \det(E_k) \det(A_{i1}) + a_{i2}(-1)^{i+2} \det(E_k) \det(A_{i2}) \\ &\quad + \cdots + a_{in}(-1)^{i+n} \det(E_k) \det(A_{in}) \\ &= \det(E_k) \det(A).\end{aligned}$$

This equation proves $\det(EA) = \det(E_k) \det(A)$ for any $n \times n$ matrix A where E_k is the corresponding elementary row operation on the $k \times k$ matrices obtained in the cofactor expansion.

The proof of the inductive step will be finished if we show that $\det(E_k) = \det(E)$. This equality follows if we let $A = I_n$ in $\det(EA) = \det(E_k) \det(A)$. Therefore, $\det(E)$ is equal to r , or 1, or -1, depending on the type of the elementary row operation E since the same is true of $\det(E_k)$ by inductive hypothesis.

Therefore, by the principle of induction, the claim is true for every $n \geq 2$. ■

As a corollary of this theorem, we can prove the multiplicativity of determinants:

Theorem 21.4. *Let A and B be $n \times n$ matrices. Then*

$$\det(AB) = \det(A) \det(B).$$

Proof. If A is non-invertible, then AB is also non-invertible and both $\det(A)$ and $\det(AB)$ are 0, proving the equality in this case.

Suppose now that A is invertible. By the Invertible Matrix Theorem, we know that A is row equivalent to I_n . Expressed in terms of elementary matrices, this means that there are elementary matrices E_1, E_2, \dots, E_ℓ such that

$$A = E_1 E_2 \cdots E_\ell I_n = E_1 E_2 \cdots E_\ell. \quad (21.2)$$

Therefore, repeatedly applying Theorem 21.3, we find that

$$\det(A) = \det(E_1) \det(E_2) \cdots \det(E_\ell). \quad (21.3)$$

If we multiply equation (21.2) by B on the right, we obtain

$$AB = E_1 E_2 \cdots E_\ell B.$$

Again, by repeatedly applying Theorem 21.3 with this product of matrices, we find

$$\det(AB) = \det(E_1 E_2 \cdots E_\ell B) = \det(E_1) \det(E_2) \cdots \det(E_\ell) \det(B).$$

From equation (21.3), the product of $\det(E_i)$'s equals $\det(A)$, so

$$\det(AB) = \det(A) \det(B)$$

which finishes the proof of the theorem. ■

We can use the multiplicative property of the determinant and the determinants of elementary matrices to calculate the determinant of a matrix in a more efficient way than using the cofactor expansion. The next activity provides an example.



Activity 21.2. Let $A = \begin{bmatrix} 1 & 1 & 2 \\ 2 & 2 & 6 \\ -1 & 2 & 1 \end{bmatrix}$.

- Use elementary row operations to reduce A to a row echelon form. Keep track of the elementary row operation you use.
- Taking into account how elementary row operations affect the determinant, use the row echelon form of A to calculate $\det(A)$.

Your work in Activity 21.2 provides an efficient method for calculating the determinant. If A is a square matrix, we use row operations given by elementary matrices E_1, E_2, \dots, E_k to row reduce A to row echelon form R . That is

$$R = E_k E_{k-1} \cdots E_2 E_1 A.$$

We know $\det(E_i)$ for each i , and since R is a triangular matrix we can find its determinant. Then

$$\det(A) = \det(E_1)^{-1} \det(E_2)^{-1} \cdots \det(E_k)^{-1} \det(R).$$

In other words, if we keep track of how the row operations affect the determinant, we can calculate the determinant of a matrix A using row operations.

Activity 21.3. Theorems 21.3 and 21.4 can be used to prove the following (part c of Theorem 16.2) that A is invertible if and only if $\det(A) \neq 0$. We see how in this activity. Let A be an $n \times n$ matrix. We can row reduce A to its reduced row echelon form R by elementary matrices E_1, E_2, \dots, E_k so that

$$R = E_1 E_2 \cdots E_k A.$$

- Suppose A is invertible. What, then, is R ? What is $\det(R)$? Can the determinant of an elementary matrix ever be 0? How do we conclude that $\det(A) \neq 0$?
- Now suppose that $\det(A) \neq 0$. What can we conclude about $\det(R)$? What, then, must R be? How do we conclude that A is invertible?

Summary: Let A be an $n \times n$ matrix. Suppose we swap rows s times and divide rows by constants k_1, k_2, \dots, k_r while computing a row echelon form $\text{REF}(A)$ of A . Then $\det(A) = (-1)^s k_1 k_2 \cdots k_r \det(\text{REF}(A))$.

Geometric Interpretation of the Determinant

Determinants have interesting and useful applications from a geometric perspective. To understand the geometric interpretation of the determinant of an $n \times n$ matrix A , we consider the image of the unit square under the transformation $T(\mathbf{x}) = A\mathbf{x}$ and see how its area changes based on A .

Activity 21.4.



- (a) Let $A = \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix}$. Start with the unit square in \mathbb{R}^2 with corners at the origin and at $(1, 1)$.

In other words, the unit square we are considering consists of all vectors $\mathbf{v} = \begin{bmatrix} x \\ y \end{bmatrix}$ where $0 \leq x \leq 1$ and $0 \leq y \leq 1$, visualized as points in the plane.

- Consider the collection of image vectors $A\mathbf{v}$ obtained by multiplying \mathbf{v} 's by A . Sketch the rectangle formed by these image vectors.
- Explain how the area of this image rectangle and the unit square is related via $\det(A)$.
- Does the relationship you found above generalize to an arbitrary $A = \begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix}$? If not, modify the relationship to hold for all diagonal matrices.

(b) Let $A = \begin{bmatrix} 2 & 1 \\ 0 & 3 \end{bmatrix}$.

- Sketch the image of the unit square under the transformation $T(\mathbf{v}) = A\mathbf{v}$. To make the sketching easier, find the images of the vectors $[0 0]^\top, [1 0]^\top, [0 1]^\top, [1 1]^\top$ as points first and then connect these images to find the image of the unit square.
- Check that the area of the parallelogram you obtained in the above part is equal to $\det(A)$.
- Does the relationship between the area and $\det(A)$ still hold if $A = \begin{bmatrix} -2 & 1 \\ 0 & 3 \end{bmatrix}$? If not, how will you modify the relationship?

It can be shown that for all 2×2 matrices a similar relationship holds.

Theorem 21.5. *For a 2×2 matrix A , the area of the image of the unit square under the transformation $T(\mathbf{x}) = A\mathbf{x}$ is equal to $|\det(A)|$. This is equivalent to saying that $|\det(A)|$ is equal to the area of the parallelogram defined by the columns of A . The area of the parallelogram is also equal to the lengths of the column vectors of A multiplied by $|\sin(\theta)|$ where θ is the angle between the two column vectors.*

There is a similar geometric interpretation of the determinant of a 3×3 matrix in terms of volume.

Theorem 21.6. *For a 3×3 matrix A , the volume of the image of the unit cube under the transformation $T(\mathbf{x}) = A\mathbf{x}$ is equal to $|\det(A)|$. This is equivalent to saying that $|\det(A)|$ is equal to the volume of the parallelepiped defined by the columns of A .*

The sign of $\det(A)$ can be interpreted in terms of the orientation of the column vectors of A . See the project in Section 16 for details.

An Explicit Formula for the Inverse and Cramer's Rule

In Section 10 we found the inverse A^{-1} using row reduction of the matrix obtained by augmenting A with I_n . However, in theoretical applications, having an explicit formula for A^{-1} can be handy.



Such an explicit formula provides us with an algebraic expression for A^{-1} in terms of the entries of A . A consequence of the formula we develop is Cramer's Rule, which can be used to provide formulas that give solutions to certain linear systems.

We begin with an interesting connection between a square matrix and the matrix of its cofactors that we explore in the next activity.

Activity 21.5. Let $A = \begin{bmatrix} 2 & 1 & 3 \\ 1 & 4 & 5 \\ 2 & -1 & 2 \end{bmatrix}$.

(a) Calculate the $(1, 1)$, $(1, 2)$, and $(1, 3)$ cofactors of A .

(b) If C_{ij} represents the (i, j) cofactor of A , then the cofactor matrix C is the matrix $C = [C_{ij}]$.

The *adjugate* matrix of A is the transpose of the cofactor matrix. In our example, the adjugate matrix of A is

$$\text{adj}(A) = \begin{bmatrix} 13 & -5 & -7 \\ 8 & -2 & -7 \\ -9 & 4 & 7 \end{bmatrix}.$$

Check the entries of this adjugate matrix with your calculations from part (a). Then calculate the matrix product

$$A \text{adj}(A).$$

(c) What do you notice about the product $A \text{adj}(A)$? How is this product related to $\det(A)$?

The result of Activity 21.5 is rather surprising, but it is valid in general. That is, if $A = [a_{ij}]$ is an invertible $n \times n$ matrix and C_{ij} is the (i, j) cofactor of A , then $A \text{adj}(A) = \det(A)I_n$. In other words, $A \left(\frac{\text{adj}(A)}{\det(A)} \right) = I_n$ and so

$$A^{-1} = \frac{1}{\det(A)} \text{adj}(A).$$

This gives us another formulation of the inverse of a matrix. To see why $A \text{adj}(A) = \det(A)I_n$, we use the row-column version of the matrix product to find the ij th entry of $A \text{adj}(A)$ as indicated by the shaded row and column

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} \begin{bmatrix} C_{11} & C_{21} & \cdots & C_{j1} & \cdots & C_{n1} \\ C_{12} & C_{22} & \cdots & C_{j2} & \cdots & C_{n2} \\ \vdots & \vdots & & \vdots & & \vdots \\ C_{1n} & C_{2n} & \cdots & C_{jn} & \cdots & C_{nn} \end{bmatrix}.$$

Thus the ij th entry of $A \text{adj}(A)$ is

$$a_{i1}C_{j1} + a_{i2}C_{j2} + \cdots + a_{in}C_{jn}. \quad (21.4)$$

Notice that if $i = j$, then expression (21.4) is the cofactor expansion of A along the i th row. So the ii th entry of $A \text{adj}(A)$ is $\det(A)$. It remains to show that the ij th entry of $A \text{adj}(A)$ is 0 when $i \neq j$.



When $i \neq j$, the expression (21.4) is the cofactor expansion of the matrix

$$\left[\begin{array}{cccc} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & \vdots & & \vdots \\ a_{j-11} & a_{j-12} & \cdots & a_{j-1n} \\ \text{shaded row} \quad a_{i1} & a_{i2} & \cdots & a_{in} \\ a_{j+11} & a_{i+12} & \cdots & a_{j+1n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{array} \right]$$

along the j th row. This matrix is the one obtained by replacing the j th row of A with the i th row of A . Since this matrix has two identical rows, it is not row equivalent to the identity matrix and is therefore not invertible. Thus, when $i \neq j$ expression (21.4) is 0. This makes $A \text{adj}(A) = \det(A)I_n$.

One consequence of the formula $A^{-1} = \frac{1}{\det(A)}\text{adj}(A)$ is Cramer's rule, which describes the solution to the equation $A\mathbf{x} = \mathbf{b}$.

Activity 21.6. Let $A = \begin{bmatrix} 3 & 1 \\ 4 & 2 \end{bmatrix}$, and let $\mathbf{b} = \begin{bmatrix} 2 \\ 6 \end{bmatrix}$.

(a) Solve the equation $A\mathbf{x} = \mathbf{b}$ using the inverse of A .

(b) Let $A_1 = \begin{bmatrix} 2 & 1 \\ 6 & 2 \end{bmatrix}$, the matrix obtained by replacing the first column of A with \mathbf{b} . Calculate $\frac{\det(A_1)}{\det(A)}$ and compare to your solution from part (a). What do you notice?

(c) Now let $A_2 = \begin{bmatrix} 3 & 2 \\ 4 & 6 \end{bmatrix}$, the matrix obtained by replacing the second column of A with \mathbf{b} . Calculate $\frac{\det(A_2)}{\det(A)}$ and compare to your solution from part (a). What do you notice?

The result from Activity 21.6 may seem a bit strange, but turns out to be true in general. The result is called *Cramer's Rule*.

Theorem 21.7 (Cramer's Rule). *Let A be an $n \times n$ invertible matrix. For any \mathbf{b} in \mathbb{R}^n , the solution \mathbf{x} of $A\mathbf{x} = \mathbf{b}$ has entries*

$$x_i = \frac{\det(A_i)}{\det(A)}$$

where A_i represents the matrix formed by replacing i th column of A with \mathbf{b} .

To see why Cramer's Rule works in general, let A be an $n \times n$ invertible matrix and $\mathbf{b} = [b_1 \ b_2 \ \cdots \ b_n]^\top$. The solution to $A\mathbf{x} = \mathbf{b}$ is

$$\mathbf{x} = A^{-1}\mathbf{b} = \frac{1}{\det(A)}\text{adj}(A)\mathbf{b} = \frac{1}{\det(A)} \begin{bmatrix} C_{11} & C_{21} & \cdots & C_{n1} \\ C_{12} & C_{22} & \cdots & C_{n2} \\ \vdots & \vdots & & \vdots \\ C_{1n} & C_{2n} & \cdots & C_{nn} \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}.$$



Expanding the product gives us

$$\mathbf{x} = \frac{1}{\det(A)} \begin{bmatrix} b_1C_{11} + b_2C_{21} + \cdots + b_nC_{n1} \\ b_1C_{12} + b_2C_{22} + \cdots + b_nC_{n2} \\ \vdots \\ b_1C_{1n} + b_2C_{2n} + \cdots + b_nC_{nn} \end{bmatrix}.$$

The expression

$$b_1C_{1j} + b_2C_{2j} + \cdots + b_nC_{nj}$$

is the cofactor expansion of the matrix

$$A_j = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1j-1} & b_1 & a_{1j+1} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2j-1} & b_2 & a_{2j+1} & \cdots & a_{2n} \\ \vdots & \vdots & & & & & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nj-1} & b_n & a_{nj+1} & \cdots & a_{nn} \end{bmatrix}$$

along the j th column, giving us the formula in Cramer's Rule.

Cramer's Rule is not a computationally efficient method. To find a solution to a linear system of n equations in n unknowns using Cramer's Rule requires calculating $n + 1$ determinants of $n \times n$ matrices – quite inefficient when n is 3 or greater. Our standard method of solving systems using Gaussian elimination is much more efficient. However, Cramer's Rule does provide a formula for the solution to $A\mathbf{x} = \mathbf{b}$ as long as A is invertible.

The Determinant of the Transpose

In this section we establish the fact that the determinant of a square matrix is the same as the determinant of its transpose.

The result is easily verified for 2×2 matrices, so we will proceed by induction and assume that the determinant of the transpose of any $(n - 1) \times (n - 1)$ matrix is the same as the determinant of its transpose. Suppose $A = [a_{ij}]$ is an $n \times n$ matrix. By definition,

$$\det(A) = a_{11}C_{11} + a_{12}C_{12} + a_{13}C_{13} + \cdots + a_{1n}C_{1n}$$

and

$$\det(A^T) = a_{11}C_{11} + a_{21}C_{21} + a_{31}C_{31} + \cdots + a_{n1}C_{n1}.$$

Note that the only terms in either determinant that contains a_{11} is $a_{11}C_{11}$. This term is the same in both determinants, so we proceed to examine other elements. Let us consider all terms in the cofactor expansion for $\det(A^T)$ that contain $a_{i1}a_{1j}$. The only summand that contains a_{i1} is $a_{i1}C_{i1}$. Letting A_{ij} be the sub-matrix of A obtained by deleting the i th row and j th column, we see that $a_{i1}C_{i1} = (-1)^{i+1}a_{i1}\det(A_{i1})$. Now let's examine the sub-matrix A_{i1} :

$$\begin{bmatrix} a_{12} & a_{13} & \cdots & a_{1j} & \cdots & a_{1n-1} & a_{1n} \\ a_{22} & a_{23} & \cdots & a_{2j} & \cdots & a_{2n-1} & a_{2n} \\ \vdots & \ddots & & \vdots & & \ddots & \\ a_{i-12} & a_{i-13} & \cdots & a_{i-1j} & \cdots & a_{i-1n-1} & a_{i-1n} \\ a_{i+12} & a_{i+13} & \cdots & a_{i+1j} & \cdots & a_{i+1n-1} & a_{i+1n} \\ a_{n2} & a_{n3} & \cdots & a_{nj} & \cdots & a_{nn-1} & a_{nn} \end{bmatrix}$$



When we expand along the first row to calculate $\det(A_{i1})$, the only term that will involve a_{1j} is

$$(-1)^{j-1+1} a_{1j} \det(A_{i1,1j}),$$

where $A_{ik,jm}$ denotes the sub-matrix of A obtained by deleting rows i and k and columns j and m from A . So the term that contains $a_{i1}a_{1j}$ in the cofactor expansion for $\det(A^T)$ is

$$(-1)^i + 1 a_{i1} (-1)^j a_{1j} \det(A_{i1,1j}) = (-1)^{i+j+1} a_{i1} a_{1j} \det(A_{i1,1j}). \quad (21.5)$$

Now we examine the cofactor expansion for $\det(A)$ to find the terms that contain $a_{i1}a_{1j}$. The quantity a_{1j} only appears in the cofactor expansion as

$$a_{1j} C_{1j} = (-1)^{1+j} a_{1j} \det(A_{1j}).$$

Now let's examine the sub-matrix A_{1j} :

$$\left[\begin{array}{ccccccc} a_{21} & a_{22} & \cdots & a_{2j-1} & a_{2j+1} & \cdots & a_{2n} \\ a_{31} & a_{32} & \cdots & a_{3j-1} & a_{3j+1} & \cdots & a_{3n} \\ \vdots & \ddots & & \vdots & \ddots & & \\ a_{i1} & a_{i2} & \cdots & a_{ij-1} & a_{ij+1} & \cdots & a_{in} \\ \vdots & & & \vdots & & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nj-1} & a_{nj+1} & \cdots & a_{nn} \end{array} \right]$$

Here is where we use the induction hypothesis. Since A_{1j} is an $(n - 1) \times (n - 1)$ matrix, its determinant can be found with a cofactor expansion down the first column. The only term in this cofactor expansion that will involve a_{i1} is

$$(-1)^{i-1+1} a_{i1} \det(A_{1i,j1}).$$

So the term that contains $a_{i1}a_{1j}$ in the cofactor expansion for $\det(A)$ is

$$(-1)^{1+j} a_{1j} (-1)^{i-1+1} a_{i1} \det(A_{1i,j1}) = (-1)^{i+j+1} a_{i1} a_{1j} \det(A_{1i,j1}). \quad (21.6)$$

Since the quantities in (21.5) and (21.6) are equal, we conclude that the terms in the two cofactor expansions are the same and

$$\det(A^T) = \det(A).$$

Row Swaps and Determinants

In this section we determine the effect of row swaps to the determinant. Let E_{rs} be the elementary matrix that swaps rows r and s in the $n \times n$ matrix $A = [a_{ij}]$. Applying E_{12} to a 2×2 matrix $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, we see that

$$\det(A) = ad - bc = -(ad - bc) = \det\left(\begin{bmatrix} c & d \\ a & b \end{bmatrix}\right) = \det(E_{12}A).$$

So swapping rows in a 2×2 matrix multiplies the determinant by -1 . Suppose that row swapping on any $(n - 1) \times (n - 1)$ matrix multiplies the determinant by -1 (in other words, we are proving our



statement by mathematical induction). Now suppose A is an $n \times n$ matrix and let $B = [b_{ij}] = E_{rs}A$. We first consider the case that $s = r + 1$ – that we swap adjacent rows. We consider two cases, $r > 1$ and $r = 1$. First let us suppose that $r > 1$. Let C_{ij} be the (i, j) cofactor of A and C'_{ij} the (i, j) cofactor of B . We have

$$\det(A) = a_{11}C_{11} + a_{12}C_{12} + \cdots + a_{1n}C_{1n}$$

and

$$\det(B) = b_{11}C'_{11} + b_{12}C'_{12} + \cdots + b_{1n}C'_{1n}.$$

Since $r > 1$, it follows that $a_{1j} = b_{1j}$ for every j . For each j the sub-matrix B_{1j} obtained from B by deleting the i th row and j th column is the same matrix as obtained from A_{ij} by swapping rows r and s . So by our induction hypothesis, we have $C'_{1j} = -C_{1j}$ for each j . Then

$$\begin{aligned}\det(B) &= b_{11}C'_{11} + b_{12}C'_{12} + \cdots + b_{1n}C'_{1n} \\ &= a_{11}(-C_{11}) + a_{12}(-C_{12}) + \cdots + a_{1n}(-C_{1n}) \\ &= -(a_{11}C_{11} + a_{12}C_{12} + \cdots + a_{1n}C_{1n}) \\ &= -\det(A).\end{aligned}$$

Now we consider the case where $r = 1$, where B is the matrix obtained from A by swapping the first and second rows. Here we will use the fact that $\det(A) = \det(A^T)$ which allows us to calculate $\det(A)$ and $\det(B)$ with the cofactor expansions down the first column. In this case we have

$$\det(A) = a_{11}C_{11} + a_{21}C_{21} + \cdots + a_{n1}C_{n1}$$

and

$$\begin{aligned}\det(B) &= b_{11}C'_{11} + b_{21}C'_{21} + \cdots + b_{n1}C'_{n1} \\ &= a_{21}C'_{11} + a_{11}C'_{21} + a_{31}C'_{31} + \cdots + a_{n1}C'_{n1}.\end{aligned}$$

For each $i \geq 3$, the sub-matrix B_{i1} is just A_{i1} with rows 1 and 2 swapped. So we have $C'_{i1} = -C_{i1}$ by our induction hypothesis. Since we swapped rows 1 and 2, we have $B_{21} = A_{11}$ and $B_{11} = A_{21}$. Thus,

$$b_{11}C'_{11} = (-1)^{1+1}b_{11}\det(A_{21}) = a_{21}\det(A_{21}) = -a_{21}C_{21}$$

and

$$b_{21}C'_{21} = (-1)^{2+1}a_{11}\det(A_{11}) = -a_{11}\det(A_{11}) = -a_{11}C_{11}.$$

Putting this all together gives us

$$\begin{aligned}\det(B) &= b_{11}C'_{11} + b_{21}C'_{21} + \cdots + b_{n1}C'_{n1} \\ &= -a_{21}C_{21} - a_{11}C_{11} + a_{31}(-C_{31}) + \cdots + a_{n1}(-C_{n1}) \\ &= -(a_{11}C_{11} + a_{21}C_{21} + \cdots + a_{n1}C_{n1}) \\ &= -\det(A).\end{aligned}$$

So we have shown that if B is obtained from A by interchanging two adjacent rows, then $\det(B) = -\det(A)$. Now we consider the general case. Suppose B is obtained from A by interchanging rows r and s , with $r < s$. We can perform this single row interchange through a sequence of adjacent row interchanges. First we swap rows r and $r + 1$, then rows $r + 1$ and $r + 2$, and continue until



we swap rows $s - 1$ and s . This places the original row r into the row s position, and the process involved $s - r$ adjacent row interchanges. Each of these interchanges multiplies the determinant by a factor of -1 . At the end of this sequence of row swaps, the original row s is now row $s - 1$. So it will take one fewer adjacent row interchanges to move this row to be row r . This sequence of $(s - r) + (s - r - 1) = 2(s - r - 1) - 1$ row interchanges produces the matrix B . Thus,

$$\det(B) = (-1)^{2(s-r)-1} \det(A) = -\det(A),$$

and interchanging any two rows multiplies the determinant by -1 .

Cofactor Expansions

We have stated that the determinant of a matrix can be calculated by using a cofactor expansion along any row or column. We use the result that swapping rows introduces a factor of -1 in the determinant to verify that result in this section. Note that in proving that $\det(A^T) = \det(A)$, we have already shown that the cofactor expansion along the first column is the same as the cofactor expansion along the first row. If we can prove that the cofactor expansion along any row is the same, then the fact that $\det(A^T) = \det(A)$ will imply that the cofactor expansion along any column is the same as well.

Now we demonstrate that the cofactor expansions along the first row and the i th row are the same. Let $A = [a_{ij}]$ be an $n \times n$ matrix. The cofactor expansion of A along the first row is

$$a_{11}C_{11} + a_{12}C_{12} + \cdots + a_{1n}C_{1n}$$

and the cofactor expansion along the i th row is

$$a_{i1}C_{i1} + a_{i2}C_{i2} + \cdots + a_{in}C_{in}.$$

Let B be the matrix obtained by swapping row i with previous rows so that row i becomes the first row and the order of the remaining rows is preserved.

$$B = \begin{bmatrix} a_{i1} & a_{i2} & \cdots & a_{ij} & \cdots & a_{in} \\ a_{11} & a_{12} & \cdots & a_{1j} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2j} & \cdots & a_{2n} \\ a_{i-11} & a_{i-12} & \cdots & a_{i-1j} & \cdots & a_{i-1n} \\ a_{i+11} & a_{i+12} & \cdots & a_{i+1j} & \cdots & a_{i+1n} \\ \vdots & \ddots & \vdots & \ddots & & \\ a_{n1} & a_{n2} & \cdots & a_{nj} & \cdots & a_{nn} \end{bmatrix}$$

Then

$$\det(B) = (-1)^{i-1} \det(A).$$

So, letting C'_{ij} be the (i, j) cofactor of B we have

$$\det(A) = (-1)^{i-1} \det(B) = (-1)^{i-1} (a_{i1}C'_{11} + a_{i2}C'_{12} + \cdots + a_{in}C'_{1n}).$$



Notice that for each j we have $B_{1j} = A_{ij}$. So

$$\begin{aligned}
 \det(A) &= (-1)^{i-1} (a_{i1}C'_{11} + a_{i2}C'_{12} + \cdots + a_{in}C'_{1n}) \\
 &= (-1)^{i-1} \left(a_{i1}(-1)^{(1+1)} \det(B_{11}) + a_{i2}(-1)^{1+2} \det(B_{12}) \right. \\
 &\quad \left. + \cdots + a_{in}(-1)^{1+n} \det(B_{1n}) \right) \\
 &= (-1)^{i-1} \left(a_{i1}(-1)^{(1+1)} \det(A_{i1}) + a_{i2}(-1)^{1+2} \det(A_{i2}) \right. \\
 &\quad \left. + \cdots + a_{in}(-1)^{1+n} \det(A_{in}) \right) \\
 &= a_{i1}(-1)^{(i+1)} \det(A_{i1}) + a_{i2}(-1)^{i+2} \det(A_{i2}) \\
 &\quad + \cdots + a_{in}(-1)^{i+n} \det(A_{in}) \\
 &= a_{i1}C_{i1} + a_{i2}C_{i2} + \cdots + a_{in}C_{in}.
 \end{aligned}$$

The LU Factorization of a Matrix

There are many instances where we have a number of systems to solve of the form $Ax = b$, all with the same coefficient matrix. The system may evolve over time so that we do not know the constant vectors b in the system all at once, but only determine them as time progresses. Each time we obtain a new vector b , we have to apply the same row operations to reduce the coefficient matrix to solve the new system. This is time repetitive and time consuming. Instead, we can keep track of the row operations in one row reduction and save ourselves a significant amount of time. One way of doing this is the *LU*-factorization (or decomposition).

To illustrate, suppose we can write the matrix A as a product $A = LU$, where

$$L = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix} \quad \text{and} \quad U = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 3 & -2 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

Let $b = [3 \ 1 \ 1 \ 3]^T$ and $x = [x_1 \ x_2 \ x_3 \ x_4]^T$, and consider the linear system $Ax = b$. If $Ax = b$, then $LUX = b$. We can solve this system without applying row operations as follows. Let $Ux = z$, where $z = [z_1 \ z_2 \ z_3 \ z_4]^T$. We can solve $Lz = b$ by using forward substitution.

The equation $Lz = b$ is equivalent to the system

$$\begin{aligned}
 z_1 &= 3 \\
 -z_1 + z_2 &= 1 \\
 z_2 + z_3 &= 1 \\
 z_4 &= 3.
 \end{aligned}$$

The first equation shows that $z_1 = 3$. Substituting into the second equation gives us $z_2 = 4$. Using this information in the third equation yields $z_3 = -3$, and then the fourth equation shows that $z_4 = 0$. To return to the original system, since $Ux = z$, we now solve this system to find the



solution vector \mathbf{x} . In this case, since U is upper triangular, we use back substitution. The equation $U\mathbf{x} = \mathbf{z}$ is equivalent to the system

$$\begin{aligned} x_1 + x_3 &= 3 \\ x_2 + 3x_3 - 2x_4 &= 4 \\ 3x_4 &= -3. \end{aligned}$$

Note that the third column of U is not a pivot column, so x_3 is a free variable. The last equation shows that $x_4 = -1$. Substituting into the second equation and solving for x_2 yields $x_2 = 2 - 3x_3$. The first equation then gives us $x_1 = 3 - x_3$. So the general solution

$$\mathbf{x} = \begin{bmatrix} 3 \\ 2 \\ 0 \\ -1 \end{bmatrix} + \begin{bmatrix} -1 \\ -3 \\ 1 \\ 0 \end{bmatrix} x_3$$

to $A\mathbf{x} = \mathbf{b}$ can be found through L and U via forward and backward substitution. If we can find a factorization of a matrix A into a lower triangular matrix L and an upper triangular matrix U , then $A = LU$ is called an *LU-factorization* or *LU-decomposition*.

We can use elementary matrices to obtain a factorization of certain matrices into products of lower triangular (the “L” in LU) and upper triangular (the “U” in LU) matrices. We illustrate with an example. Let

$$A = \begin{bmatrix} 1 & 0 & 1 & 0 \\ -1 & 1 & 2 & -2 \\ 0 & 1 & 3 & 1 \\ 1 & 0 & 1 & 0 \end{bmatrix}.$$

Our goal is to find an upper triangular matrix U and a lower triangular matrix L so that $A = LU$. We begin by row reducing A to an upper triangular matrix, keeping track of the elementary matrices used to perform the row operations. We start by replacing the entries below the $(1, 1)$ entry in A with zeros. The elementary matrices that perform these operations are

$$E_1 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad \text{and} \quad E_2 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -1 & 0 & 0 & 1 \end{bmatrix},$$

and

$$E_2 E_1 A = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 3 & -2 \\ 0 & 1 & 3 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

We next zero out the entries below the $(2, 2)$ entry as

$$E_3 E_2 E_1 A = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 3 & -2 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix},$$

where

$$E_3 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

The product $E_3E_2E_1A$ is an upper triangular matrix U . So we have

$$E_3E_2E_1A = U$$

and

$$A = E_1^{-1}E_2^{-1}E_3^{-1}U,$$

where

$$E_1^{-1}E_2^{-1}E_3^{-1} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix}$$

is a lower triangular matrix L . So we have decomposed the matrix A into a product $A = LU$, where L is lower triangular and U is upper triangular. Since every matrix is row equivalent to a matrix in row echelon form, we can always find an upper triangular matrix U in this way. However, we may not always obtain a corresponding lower triangular matrix, as the next example illustrates.

Suppose we change the problem slightly and consider the matrix

$$B = \begin{bmatrix} 1 & 0 & 1 & 0 \\ -1 & 1 & 2 & -2 \\ 0 & 1 & 3 & 1 \\ 1 & 0 & 0 & 1 \end{bmatrix}.$$

Using the same elementary matrices E_1 , E_2 , and E_3 as earlier, we have

$$E_3E_2E_1B = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 3 & -2 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & -1 & 1 \end{bmatrix}.$$

To reduce B to row-echelon form now requires a row interchange. Letting

$$E_4 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

brings us to

$$E_4E_3E_2E_1B = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 3 & -2 \\ 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & 3 \end{bmatrix}.$$



So in this case we have $U = E_4E_3E_2E_1B$, but

$$E_1^{-1}E_2^{-1}E_3^{-1}E_4^{-1} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \end{bmatrix}$$

is not lower triangular. The difference in this latter example is that we needed a row swap to obtain the upper triangular form.

Examples

What follows are worked examples that use the concepts from this section.

Example 21.8.

- (a) If A, B are $n \times n$ matrices with $\det(A) = 3$ and $\det(B) = 2$, evaluate the following determinant values. Briefly justify.

- i. $\det(A^{-1})$
- ii. $\det(ABA^T)$
- iii. $\det(A^3(BA)^{-1}(AB)^2)$

- (b) If the determinant of $\begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix}$ is m , find the determinant of each of the following matrices.

- i. $\begin{bmatrix} a & b & c \\ 2d & 2e & 2f \\ g & h & i \end{bmatrix}$
- ii. $\begin{bmatrix} d & e & f \\ g & h & i \\ a & b & c \end{bmatrix}$
- iii. $\begin{bmatrix} a & b & c \\ g-2d & h-2e & i-2f \\ a+d & b+e & c+f \end{bmatrix}$

Example Solution.

- (a) Assume that $\det(A) = 3$ and $\det(B) = 2$.

- i. Since $\det(A) \neq 0$, we know that A is invertible. Since $1 = \det(I_n) = \det(AA^{-1}) = \det(A)\det(A^{-1})$, it follows that $\det(A^{-1}) = \frac{1}{\det(A)} = \frac{1}{3}$.



ii. We know that $\det(A^T) = \det(A)$, so

$$\begin{aligned}\det(ABA^T) &= \det(A)\det(B)\det(A^T) \\ &= \det(A)\det(B)\det(A) \\ &= (3)(2)(3) \\ &= 18.\end{aligned}$$

iii. Using properties of determinants gives us

$$\begin{aligned}\det(A^3(BA)^{-1}(AB)^2) &= \det(A^3)\det((BA)^{-1})\det((AB)^2) \\ &= (\det(A))^3 \left(\frac{1}{\det(AB)} \right) (\det(AB))^2 \\ &= 27 \left(\frac{1}{\det(A)\det(B)} \right) (\det(A)\det(B))^2 \\ &= \frac{(27)(6^2)}{6} \\ &= 162.\end{aligned}$$

(b) Assume that $\det \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix} = m$.

i. Multiplying a row by a scalar multiples the determinant by that scalar, so

$$\det \begin{pmatrix} a & b & c \\ 2d & 2e & 2f \\ g & h & i \end{pmatrix} = 2m.$$

ii. Interchanging two rows multiples the determinant by -1 . It takes two row swaps in the original matrix to obtain this one, so

$$\det \begin{pmatrix} d & e & f \\ g & h & i \\ a & b & c \end{pmatrix} = (-1)^2 m = m.$$

iii. Adding a multiple of a row to another does not change the determinant of the matrix. Since there is a row swap needed to get this matrix from the original we have

$$\det \begin{pmatrix} a & b & c \\ g - 2d & h - 2e & i - 2f \\ a + d & b + e & c + f \end{pmatrix} = -m.$$

Example 21.9. Let $A = \begin{bmatrix} 2 & 8 & 0 \\ 2 & 2 & -3 \\ 1 & 2 & 7 \end{bmatrix}$.

(a) Find an LU factorization for A .



- (b) Use the LU factorization with forward substitution and back substitution to solve the system $A\mathbf{x} = [18 \ 3 \ 12]^\top$.

Example Solution.

- (a) We row reduce A to an upper triangular matrix by applying elementary matrices. First notice that if $E_1 = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$, then

$$E_1 A = \begin{bmatrix} 2 & 8 & 0 \\ 0 & -6 & -3 \\ 1 & 2 & 7 \end{bmatrix}.$$

Letting $E_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -\frac{1}{2} & 0 & 1 \end{bmatrix}$ gives us

$$E_2 E_1 A = \begin{bmatrix} 2 & 8 & 0 \\ 0 & -6 & -3 \\ 0 & -2 & 7 \end{bmatrix}.$$

Finally, when $E_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -\frac{1}{3} & 1 \end{bmatrix}$ we have

$$U = E_3 E_2 E_1 A = \begin{bmatrix} 2 & 8 & 0 \\ 0 & -6 & -3 \\ 0 & 0 & 8 \end{bmatrix}.$$

This gives us $E_3 E_2 E_1 A = U$, so we can take

$$L = E_1^{-1} E_2^{-1} E_3^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ \frac{1}{2} & \frac{1}{3} & 1 \end{bmatrix}.$$

- (b) To solve the system $A\mathbf{x} = \mathbf{b}$, where $\mathbf{b} = [18 \ 3 \ 12]^\top$, we use the LU factorization of A and solve $L\mathbf{U}\mathbf{x} = \mathbf{b}$. Let $\mathbf{x} = [x_1 \ x_2 \ x_3]^\top$ and let $\mathbf{z} = [z_1 \ z_2 \ z_3]^\top$ with $U\mathbf{x} = \mathbf{z}$ so that $L\mathbf{z} = L(U\mathbf{x}) = A\mathbf{x} = \mathbf{b}$. First we solve $L\mathbf{z} = [18 \ 3 \ 12]^\top$ to find \mathbf{z} using forward substitution. The first row of L shows that $z_1 = 18$ and the second row that $z_1 + z_2 = 3$. So $z_2 = -15$. The third row of L gives us $\frac{1}{2}z_1 + \frac{1}{3}z_2 + z_3 = 12$, so $z_3 = 12 - 9 + 5 = 8$. Now to find \mathbf{x} we solve $U\mathbf{x} = \mathbf{z}$ using back substitution. The third row of U tells us that $8x_3 = 8$ or that $x_3 = 1$. The second row of U shows that $-6x_2 - 3x_3 = -15$ or $x_2 = 2$. Finally, the first row of U gives us $2x_1 + 8x_2 = 18$, or $x_1 = 1$. So the solution to $A\mathbf{x} = \mathbf{b}$ is $\mathbf{x} = [1 \ 2 \ 1]^\top$.

Summary

- The elementary row operations have the following effects on the determinant:
 - If we multiply a row of a matrix by a constant k , then the determinant is multiplied by k .
 - If we swap two rows of a matrix, then the determinant changes sign.
 - If we add a multiple of a row of a matrix to another, the determinant does not change.
- Each of the elementary row operations can be achieved by multiplication by elementary matrices. To obtain the elementary matrix corresponding to an elementary row operation, we perform the operation on the identity matrix.
- Let A be an $n \times n$ invertible matrix. For any \mathbf{b} in \mathbb{R}^n , the solution \mathbf{x} of $A\mathbf{x} = \mathbf{b}$ has entries

$$x_i = \frac{\det(A_i(\mathbf{b}))}{\det(A)}$$

where $A_i(\mathbf{b})$ represents the matrix formed by replacing i th column of A with \mathbf{b} .

- Let A be an invertible $n \times n$ matrix. Then

$$A^{-1} = \frac{1}{\det(A)} \text{adj } A$$

where the $\text{adj } A$ matrix, the *adjugate of A* , is defined as the matrix whose ij -th entry is C_{ji} , the ji -th cofactor of A .

- For a 2×2 matrix A , the area of the image of the unit square under the transformation $T(\mathbf{x}) = A\mathbf{x}$ is equal to $|\det(A)|$, which is also equal to the area of the parallelogram defined by the columns of A .
- For a 3×3 matrix A , the volume of the image of the unit cube under the transformation $T(\mathbf{x}) = A\mathbf{x}$ is equal to $|\det(A)|$, which is also equal to the volume of the parallelepiped defined by the columns of A .

Exercises

- (1) Find a formula for $\det(rA)$ in terms of r and $\det(A)$, where A is an $n \times n$ matrix and r is a scalar. Explain why your formula is valid.
- (2) Find $\det(A)$ by hand using elementary row operations where

$$A = \begin{bmatrix} 1 & 2 & -1 & 3 \\ -1 & -2 & 3 & -1 \\ -2 & -1 & 2 & -3 \\ 1 & 8 & -3 & 8 \end{bmatrix}.$$



(3) Consider the matrix $A = \begin{bmatrix} 4 & -1 & -1 & -1 \\ -1 & 4 & -1 & -1 \\ -1 & -1 & 4 & -1 \\ -1 & -1 & -1 & 4 \end{bmatrix}$. We will find $\det(A)$ using elementary

row operations. (This matrix arises in graph theory, and its determinant gives the number of spanning trees in the complete graph with 5 vertices. This number is also equal to the number of labeled trees with 5 vertices.)

- (a) Add rows R_2 , R_3 and R_4 to the first row in that order.
- (b) Then add the new R_1 to rows R_2 , R_3 and R_4 to get a triangular matrix B .
- (c) Find the determinant of B . Then use $\det(B)$ and properties of how elementary row operations affect determinants to find $\det(A)$.
- (d) Generalize your work to find the determinant of the $n \times n$ matrix

$$A = \begin{bmatrix} n & -1 & -1 & \cdots & -1 & -1 \\ -1 & n & -1 & \cdots & -1 & -1 \\ \vdots & \vdots & \vdots & \cdots & \vdots & \vdots \\ -1 & -1 & -1 & \cdots & -1 & n \end{bmatrix}.$$

(4) For which matrices A , if any, is $\det(A) = -\det(-A)$? Justify your answer.

(5) Find the inverse A^{-1} of $A = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 2 & 0 & 1 \end{bmatrix}$ using the adjugate matrix.

(6) For an invertible $n \times n$ matrix A , what is the relationship between $\det(A)$ and $\det(\text{adj } A)$? Justify your result.

(7) Let $A = \begin{bmatrix} a & b & 1 \\ c & d & 2 \\ e & f & 3 \end{bmatrix}$, and assume that $\det(A) = 2$. Determine the determinants of each of the following.

(a) $B = \begin{bmatrix} a & b & 1 \\ 3c & 3d & 6 \\ e+a & f+b & 4 \end{bmatrix}$

(b) $C = \begin{bmatrix} 2e & 2f & 6 \\ 2c-2e & 2d-2f & -2 \\ 2a & 2b & 2 \end{bmatrix}$

(8) Find the area of the parallelogram with one vertex at the origin and adjacent vertices at $(1, 2)$ and (a, b) . For which (a, b) is the area 0? When does this happen geometrically?

(9) Find the volume of the parallelepiped with one vertex at the origin and three adjacent vertices at $(3, 2, 0)$, $(1, 1, 1)$ and $(1, 3, c)$ where c is unknown. For which c , is the volume 0? When does this happen geometrically?

(10) Label each of the following statements as True or False. Provide justification for your response.



- (a) **True/False** If two rows are equal in A , then $\det(A) = 0$.
- (b) **True/False** If A is a square matrix and R is a row echelon form of A , then $\det(A) = \det(R)$.
- (c) **True/False** If a matrix A is invertible, then 0 is not an eigenvalue of A .
- (d) **True/False** If A is a 2×2 matrix for which the image of the unit square under the transformation $T(\mathbf{x}) = A\mathbf{x}$ has zero area, then A is non-invertible.
- (e) **True/False** Row operations do not change the determinant of a square matrix.
- (f) **True/False** If A_{ij} is the matrix obtained from a square matrix $A = [a_{ij}]$ by deleting the i th row and j th column of A , then

$$\begin{aligned} & a_{i1}(-1)^{i+1} \det(A_{i1}) + a_{i2}(-1)^{i+2} \det(A_{i2}) + \dots \\ & \quad + a_{in}(-1)^{i+n} \det(A_{in}) \\ &= a_{1j}(-1)^{j+1} \det(A_{1j}) + a_{2j}(-1)^{j+2} \det(A_{2j}) + \dots \\ & \quad + a_{nj}(-1)^{j+n} \det(A_{nj}) \end{aligned}$$

for any i and j between 1 and n .

- (g) **True/False** If A is an invertible matrix, then $\det(A^T A) > 0$.



Part V

Vector Spaces

Section 22

Vector Spaces

Focus Questions

By the end of this section, you should be able to give precise and thorough answers to the questions listed below. You may want to keep these questions in mind to focus your thoughts as you complete the section.

- What is a vector space?
- What is a subspace of a vector space?
- What is a linear combination of vectors in a vector space V ?
- What is the span of a set of vectors in a vector space V ?
- What special structure does the span of a set of vectors in a vector space V have?
- Why is the vector space concept important?

Application: The Hat Puzzle

In a New York Times article (April 10, 2001) “Why Mathematicians Now Care About Their Hat Color”, the following puzzle is posed.

“Three players enter a room and a red or blue hat is placed on each person’s head. The color of each hat is determined by a coin toss, with the outcome of one coin toss having no effect on the others. Each person can see the other players’ hats but not his own.

No communication of any sort is allowed, except for an initial strategy session before the game begins. Once they have had a chance to look at the other hats, the players must simultaneously guess the color of their own hats or pass. The group shares a hypothetical \$3 million prize if at least one player guesses correctly and no players guess incorrectly.”

The game can be played with more players, and the problem is to find a strategy for the group that maximizes its chance of winning. One strategy is for a designated player to make a random guess and for the others to pass. This gives a 50% chance of winning. However, there are much better strategies that provide a nearly 100% probability of winning as the number of players increases. One such strategy is based on Hamming codes and subspaces of a particular vector space to implement the most effective approach.

Introduction

We have previously seen that \mathbb{R}^n and the set of fixed size matrices have a nice algebraic structure when endowed with the addition and scalar multiplication operations. In fact, as we will see, there are many other sets of elements that have the same kind of structure with natural addition and scalar multiplication operations. Due to this underlying similar structure, these sets are connected in some way and can all be studied jointly. Mathematicians look for these kinds of connections between seemingly dissimilar objects and, from a mathematical standpoint, it is convenient to study all of these similar structures at once by combining them into a larger collection. This motivates the idea of a *vector space* that we will investigate in this chapter.

An example of a set that has a structure similar to vectors is a collection of polynomials. Let \mathbb{P}_1 be the collection of all polynomials of degree less than or equal to 1 with real coefficients. That is,

$$\mathbb{P}_1 = \{a_0 + a_1 t : a_0, a_1 \in \mathbb{R}\}.$$

So, for example, the polynomials $2 + t$, $5t$, -7 , and $\sqrt{12} - \pi t$ are in \mathbb{P}_1 , but \sqrt{t} is not in \mathbb{P}_1 . Two polynomials $a(t) = a_0 + a_1 t$ and $b(t) = b_0 + b_1 t$ in \mathbb{P}_1 are equal if $a_0 = b_0$ and $a_1 = b_1$.

We define addition of polynomials in \mathbb{P}_1 by adding the coefficients of the like degree terms. So if $a(t) = a_0 + a_1 t$ and $b(t) = b_0 + b_1 t$, then the polynomial sum of $a(t)$ and $b(t)$ is

$$a(t) + b(t) = (a_0 + a_1 t) + (b_0 + b_1 t) = (a_0 + b_0) + (a_1 + b_1)t.$$

So, for example,

$$(2 + 3t) + (-1 + 5t) = (2 + (-1)) + (3 + 5)t = 1 + 8t.$$

We now consider the properties of the addition operation. For example, we can ask if polynomial addition is commutative. That is, if $a(t)$ and $b(t)$ are in \mathbb{P}_1 , must it be the case that

$$a(t) + b(t) = b(t) + a(t)?$$

To show that addition is commutative in \mathbb{P}_1 , we choose arbitrary polynomials $a(t) = a_0 + a_1 t$ and $b(t) = b_0 + b_1 t$ in \mathbb{P}_1 . Then we have

$$\begin{aligned} a(t) + b(t) &= (a_0 + b_0) + (a_1 + b_1)t \\ &= (b_0 + a_0) + (b_1 + a_1)t \\ &= b(t) + a(t). \end{aligned}$$

Note that in the middle step, we used the definition of equality of polynomials since $a_0 + b_0 = b_0 + a_0$ and $a_1 + b_1 = b_1 + a_1$ due to the fact that addition of real numbers is commutative. So addition of elements in \mathbb{P}_1 is a commutative operation.



Preview Activity 22.1.

(1) Now we investigate other properties of addition in \mathbb{P}_1 .

- (a) To show addition is associative in \mathbb{P}_1 , we need to verify that if $a(t) = a_0 + a_1t$, $b(t) = b_0 + b_1t$, and $c(t) = c_0 + c_1t$ are in \mathbb{P}_1 , it must be the case that

$$(a(t) + b(t)) + c(t) = a(t) + (b(t) + c(t)).$$

Either verify this property by using the definition of two polynomials being equal, or give a counterexample to show the equality fails in that case.

- (b) Find a polynomial $z(t) \in \mathbb{P}_1$ such that

$$a(t) + z(t) = a(t)$$

for all $a(t) \in \mathbb{P}_1$. This polynomial is called the *zero* polynomial or the *additive identity* polynomial in \mathbb{P}_1 .

- (c) If $a(t) = a_0 + a_1t$ is an element of \mathbb{P}_1 , is there an element $p(t) \in \mathbb{P}_1$ such that

$$a(t) + p(t) = z(t),$$

where $z(t)$ is the additive identity polynomial you found above? If not, why not? If so, what polynomial is $p(t)$? Explain.

(2) We can also define a multiplication of polynomials by scalars (real numbers).

- (a) What element in \mathbb{P}_1 could be the scalar multiple $\frac{1}{2}(2 + 3t)$?
- (b) In general, if k is a scalar and $a(t) = a_0 + a_1t$ is in \mathbb{P}_1 , how do we define the scalar multiple $ka(t)$ in \mathbb{P}_1 ?
- (c) If k is a scalar and $a(t) = a_0 + a_1t$ and $b(t) = b_0 + b_1t$ are elements in \mathbb{P}_1 , is it true that

$$k(a(t) + b(t)) = ka(t) + kb(t)?$$

If no, explain why. If yes, verify your answer using the definition of two polynomials being equal.

- (d) If k and m are scalars and $a(t) = a_0 + a_1t$ is an element in \mathbb{P}_1 , is it true that

$$(k + m)a(t) = ka(t) + ma(t)?$$

If no, explain why. If yes, verify your answer.

- (e) If k and m are scalars and $a(t) = a_0 + a_1t$ is an element in \mathbb{P}_1 , is it true that

$$(km)a(t) = k(ma(t))?$$

If no, explain why. If yes, verify your answer.

- (f) If $a(t) = a_0 + a_1t$ is an element of \mathbb{P}_1 , is it true that

$$1a(t) = a(t)?$$

If no, explain why. If yes, verify your answer.



Spaces with Similar Structure to \mathbb{R}^n

Mathematicians look for patterns and for similarities between mathematical objects. In doing so, mathematicians often consider larger collections of objects that are sorted according to their similarities and then study these collections rather than just the objects themselves. This perspective can be very powerful – whatever can be shown to be true about an arbitrary element in a collection will then be true for every specific element in the collection. In this section we study the larger collection of sets that share the algebraic structure of vectors in \mathbb{R}^n . These sets are called *vector spaces*.

In Preview Activity 22.1, we showed that the set \mathbb{P}_1 of polynomials of degree less than or equal to one with real coefficients, with the operations of addition and scalar multiplication defined by

$$(a_0 + a_1t) + (b_0 + b_1t) = (a_0 + b_0) + (a_1 + b_1)t \quad \text{and} \quad k(a_0 + a_1t) = (ka_0) + (ka_1)t,$$

has a structure similar to \mathbb{R}^2 .

By structure we mean how the elements in the set relate to each other under addition and multiplication by scalars. That is, if $a(t) = a_0 + a_1t$, $b(t)$, and $c(t)$ are elements of \mathbb{P}_1 and k and m are scalars, then

- (1) $a(t) + b(t)$ is an element of \mathbb{P}_1 ,
- (2) $a(t) + b(t) = b(t) + a(t)$,
- (3) $(a(t) + b(t)) + c(t) = a(t) + (b(t) + c(t))$,
- (4) there is a zero polynomial $z(t)$ (namely, $0 + 0t$) in \mathbb{P}_1 so that $a(t) + z(t) = a(t)$,
- (5) there is an element $-a(t)$ in \mathbb{P}_1 (namely, $(-a_0) + (-a_1)t$) so that $a(t) + (-a(t)) = z(t)$,
- (6) $ka(t)$ is an element of \mathbb{P}_1 ,
- (7) $(k + m)a(t) = ka(t) + ma(t)$,
- (8) $k(a(t) + b(t)) = ka(t) + kb(t)$,
- (9) $(km)a(t) = k(ma(t))$,
- (10) $1a(t) = a(t)$.

The properties we saw for polynomials in \mathbb{P}_1 stated above are the same as the properties for vector addition and multiplication by scalars in \mathbb{R}^n , as well as matrix addition and multiplication by scalars identified in Section 8. This indicates that polynomials in \mathbb{P}_1 , vectors in \mathbb{R}^n , and the set of $m \times n$ matrices behave in much the same way as regards their addition and multiplication by scalars. There is an even closer connection between linear polynomials and vectors in \mathbb{R}^2 . An element $a(t) = a_0 + a_1t$ in \mathbb{P}_1 can be naturally associated with the vector $\begin{bmatrix} a_0 \\ a_1 \end{bmatrix}$ in \mathbb{R}^2 . All the results of polynomial addition and multiplication by scalars then translate to corresponding results of addition and multiplication by scalars of vectors in \mathbb{R}^2 . So for all intents and purposes, as far as addition and multiplication by scalars is concerned, there is no difference between elements in \mathbb{P}_1



and vectors in \mathbb{R}^2 – the only difference is how we choose to present the elements (as polynomials or as vectors). This sameness of structure of our sets as it relates to addition and multiplication by scalars is the type of similarity mentioned in the introduction. We can study all of the types of objects that exhibit this same structure at one time by studying vector spaces.

Vector Spaces

We defined vector spaces in the context of subspaces of \mathbb{R}^n in Definition 12.1. In general, any set that has the same kind of additive and multiplicative structure as our sets of vectors, matrices, and linear polynomials is called a vector space. As we will see, the ideas that we introduced about subspaces of \mathbb{R}^n apply to vector spaces in general, so the material in this chapter should have a familiar feel.

Definition 22.1. A set V on which an operation of addition and a multiplication by scalars is defined is a **vector space** if for all \mathbf{u} , \mathbf{v} , and \mathbf{w} in V and all scalars a and b :

- (1) $\mathbf{u} + \mathbf{v}$ is an element of V (we say that V is *closed* under the addition in V),
- (2) $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$ (we say that the addition in V is *commutative*),
- (3) $(\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w})$ (we say that the addition in V is *associative*),
- (4) there is a zero vector $\mathbf{0}$ in V so that $\mathbf{u} + \mathbf{0} = \mathbf{u}$ (we say that V contains an *additive identity* $\mathbf{0}$),
- (5) for each \mathbf{x} in V there is an element \mathbf{y} in V so that $\mathbf{x} + \mathbf{y} = \mathbf{0}$ (we say that V contains an *additive inverse* \mathbf{y} for each element \mathbf{x} in V),
- (6) $a\mathbf{u}$ is an element of V (we say that V is *closed* under multiplication by scalars),
- (7) $(a + b)\mathbf{u} = a\mathbf{u} + b\mathbf{u}$ (we say that *multiplication by scalars distributes over scalar addition*),
- (8) $a(\mathbf{u} + \mathbf{v}) = a\mathbf{u} + a\mathbf{v}$ (we say that *multiplication by scalars distributes over addition in V*),
- (9) $(ab)\mathbf{u} = a(b\mathbf{u})$,
- (10) $1\mathbf{u} = \mathbf{u}$.

Note. Unless otherwise stated, in this book the scalars will refer to real numbers. However, we can define vector spaces where scalars are complex numbers, or rational numbers, or integers modulo p where p is a prime number, or, more generally, elements of a field. A field is an algebraic structure which generalizes the structure of real numbers and rational numbers under the addition and multiplication operations. Since we will focus on the real numbers as scalars, the reader is not required to be familiar with the concept of a field.

Because of the similarity of the way elements in vector spaces behave compared to vectors in \mathbb{R}^n , we call the elements in a vector space *vectors*. There are many examples of vectors spaces, which is what makes this idea so powerful.

Example 22.2.



- (1) The space \mathbb{R}^n of all vectors with n components is a vector space using the standard vector addition and multiplication by scalars. The zero element is the zero vector $\mathbf{0}$ whose components are all 0.
- (2) The set \mathbb{P}_1 of all polynomials of degree less than or equal to 1 with addition and scalar multiplication as defined earlier. Recall that \mathbb{P}_1 is essentially the same as \mathbb{R}^2 .
- (3) The properties listed in the introduction for \mathbb{P}_1 are equally true for the collection of all polynomials of degree less than or equal to some fixed number. We label as \mathbb{P}_n this set of all polynomials of degree less than or equal to n , with the standard addition and scalar multiplication. Note that \mathbb{P}_n is essentially the same as \mathbb{R}^{n+1} . More generally, the space \mathbb{P} of all polynomials is also a vector space with standard addition and scalar multiplication.
- (4) As a subspace of \mathbb{R}^n , the eigenspace of an $n \times n$ matrix corresponding to an eigenvalue λ is a vector space.
- (5) As a subspace of \mathbb{R}^n , the null space of an $m \times n$ matrix is a vector space.
- (6) As a subspace of \mathbb{R}^m , the column space of an $m \times n$ matrix is a vector space.
- (7) The span of a set of vectors in \mathbb{R}^n is a subspace of \mathbb{R}^n , and is therefore a vector space.
- (8) Let V be a vector space and let $\mathbf{0}$ be the additive identity in V . The set $\{\mathbf{0}\}$ is a vector space in which $\mathbf{0} + \mathbf{0} = \mathbf{0}$ and $k\mathbf{0} = \mathbf{0}$ for any scalar k . This space is called the *trivial* vector space.
- (9) The space $\mathcal{M}_{m \times n}$ (or $\mathcal{M}_{m \times n}(\mathbb{R})$) when it is important to indicate that the entries of our matrices are real numbers) of all $m \times n$ matrices with real entries with the standard addition and multiplication by scalars we have already defined. In this case, $\mathcal{M}_{m \times n}$ is essentially the same vector space as \mathbb{R}^{mn} .
- (10) The space \mathcal{F} of all functions from \mathbb{R} to \mathbb{R} , where we define the sum of two functions f and g in \mathcal{F} as the function $f + g$ satisfying

$$(f + g)(x) = f(x) + g(x)$$

for all real numbers x , and the scalar multiple cf of the function f by the scalar c to be the function satisfying

$$(cf)(x) = cf(x)$$

for all real numbers x . The verification of the vector space properties for this space is left to the reader.

- (11) The space \mathbb{R}^∞ of all infinite real sequences (x_1, x_2, x_3, \dots) . We define addition and scalar multiplication termwise:

$$(x_1, x_2, x_3, \dots) + (y_1, y_2, y_3, \dots) = (x_1 + y_1, x_2 + y_2, x_3 + y_3, \dots),$$

$$c(x_1, x_2, x_3, \dots) = (cx_1, cx_2, cx_3, \dots)$$

is a vector space. In addition, the set of convergent sequences inside \mathbb{R}^∞ forms a vector space using this addition and multiplication by scalars (as we did in \mathbb{R}^n , we will call this set of convergent sequences a subspace of \mathbb{R}^∞).



- (12) (For those readers who are familiar with differential equations). The set of solutions to a second order homogeneous differential equation forms a vector space under addition and scalar multiplication defined as in the space \mathcal{F} above.
- (13) The set of polynomials of positive degree in \mathbb{P}_1 is not a vector space using the standard addition and multiplication by scalars in \mathbb{P}_1 is not a vector space. Notice that $t + (-t)$ is not a polynomial of positive degree, and so this set is not closed under addition.
- (14) The color space where each color is assigned an RGB (red, green, blue) coordinate between 0 and 255, with addition and scalar multiplication defined component-wise, however, does not define a vector space. The color space is not closed under either operation due to the color coordinates being integers ranging from 0 to 255.

It is important to note that the set of defining properties of a vector space is intended to be a minimum set. Any other properties of a vector space must be verified or proved using the defining properties. For example, in \mathbb{R}^n it is clear that the scalar multiple $0\mathbf{v}$ is the zero vector for any vector \mathbf{v} in \mathbb{R}^n . This might be true in any vector space, but it is not a defining property. Therefore, if this property is true, then we must be able to prove it using just the defining properties. To see how this might work, let \mathbf{v} be any vector in a vector space V . We want to show that $0\mathbf{v} = \mathbf{0}$ (the existence of the zero vector is property (4)). Using the fact that $0 + 0 = 0$ and that scalar multiplication distributes over scalar addition, we can see that

$$0\mathbf{v} = (0 + 0)\mathbf{v} = 0\mathbf{v} + 0\mathbf{v}.$$

Property (5) tells us that V contains an additive inverse for every vector in V , so let \mathbf{u} be an additive inverse of the vector $0\mathbf{v}$ in V . Then $0\mathbf{v} + \mathbf{u} = \mathbf{0}$ ¹ and so

$$\begin{aligned} 0\mathbf{v} + \mathbf{u} &= (0\mathbf{v} + 0\mathbf{v}) + \mathbf{u} \\ \mathbf{0} &= 0\mathbf{v} + (0\mathbf{v} + \mathbf{u}) \\ \mathbf{0} &= 0\mathbf{v} + \mathbf{0}. \end{aligned}$$

Now $\mathbf{0}$ has the property that $\mathbf{0} + \mathbf{w} = \mathbf{w} + \mathbf{0} = \mathbf{w}$ for any vector \mathbf{w} in V (by properties (4) and (2)), and so we can conclude that

$$\mathbf{0} = 0\mathbf{v}.$$

Activity 22.1. Another property that will be useful is a cancellation property. In the set of real numbers we know that if $a + b = c + b$, then $a = c$, and we verify this by subtracting b from both sides. This is the same as adding the additive inverse of b to both sides, so we ought to be able to make the same argument using additive inverses in a vector space. To see how, let \mathbf{u} , \mathbf{v} , and \mathbf{w} be vectors in a vector space and suppose that

$$\mathbf{u} + \mathbf{w} = \mathbf{v} + \mathbf{w}. \tag{22.1}$$

- (a) Why does our space contain an additive inverse \mathbf{z} of \mathbf{w} ?

¹It is very important to keep track of the different kinds of zeros here – the boldface zero $\mathbf{0}$ is the additive identity in the vector space and the non-bold 0 is the scalar zero.

(b) Now add the vector \mathbf{z} to both sides of equation (22.1) to obtain

$$(\mathbf{u} + \mathbf{w}) + \mathbf{z} = (\mathbf{v} + \mathbf{w}) + \mathbf{z}. \quad (22.2)$$

Which property of a vector space allows us to state the following equality?

$$\mathbf{u} + (\mathbf{w} + \mathbf{z}) = \mathbf{v} + (\mathbf{w} + \mathbf{z}). \quad (22.3)$$

(c) Now use the properties of additive inverses and the additive identity to explain why $\mathbf{u} = \mathbf{v}$. Conclude that we have a cancellation law for addition in any vector space.

We should also note that the definition of a vector space only states the existence of a zero vector and an additive inverse for each vector in the space, and does not say that there cannot be more than one zero vector or more than one additive inverse of a vector in the space. The reason why is that the uniqueness of the zero vector and an additive inverse of a vector can be proved from the defining properties of a vector space, and so we don't list this consequence as a defining property. Similarly, the defining properties of a vector space do not state that the additive inverse of a vector \mathbf{v} is the scalar multiple $(-1)\mathbf{v}$. Verification of these properties are left for the exercises. We summarize the results of this section in the following theorem.

Theorem 22.3. *Let V be any vector space with identity $\mathbf{0}$.*

- $0\mathbf{v} = \mathbf{0}$ for any vector \mathbf{v} in V .
- The vector $\mathbf{0}$ is unique.
- $c\mathbf{0} = \mathbf{0}$ for any scalar c .
- For any \mathbf{v} in V , the additive inverse of \mathbf{v} is unique.
- The additive inverse of a vector \mathbf{v} in V is the vector $(-1)\mathbf{v}$.
- If \mathbf{u} , \mathbf{v} , and \mathbf{w} are in V and $\mathbf{u} + \mathbf{w} = \mathbf{v} + \mathbf{w}$, then $\mathbf{u} = \mathbf{v}$.

Subspaces

In Section 12 we saw that \mathbb{R}^n contained subsets that we called subspaces that had the same algebraic structure as \mathbb{R}^n . The same idea applies to vector spaces in general.

Activity 22.2. Let $H = \{at : a \in \mathbb{R}\}$. Notice that H is a subset of \mathbb{P}_1 .

- (a) Is H closed under the addition in \mathbb{P}_1 ? Verify your answer.
- (b) Does H contain the zero vector from \mathbb{P}_1 ? Verify your answer.
- (c) Is H closed under multiplication by scalars? Verify your answer.
- (d) Explain why H satisfies every other property of the definition of a vector space automatically just by being a subset of \mathbb{P}_1 and using the same operations as in \mathbb{P}_1 . Conclude that H is a vector space.



Activity 22.2 illustrates an important point. There is a fundamental difference in the types of properties that define a vector space. Some of the properties that define a vector space are true for any subset of the vector space because they are properties of the operations (such as the commutative and associative properties). The other properties (closure, the inclusion of the zero vector, and the inclusion of additive inverses) are set properties, not properties of the operations. So these three properties have to be specifically checked to see if a subset of a vector space is also a vector space. This leads to the definition of a *subspace*, a subset of a vector space which is a vector space itself.

Definition 22.4. A subset H of a vector space V is a **subspace** of V if

- (1) whenever \mathbf{u} and \mathbf{v} are in H it is also true that $\mathbf{u} + \mathbf{v}$ is in H (that is, H is *closed* under addition),
- (2) whenever \mathbf{u} is in H and a is a scalar it is also true that $a\mathbf{u}$ is in H (that is, H is *closed* under scalar multiplication),
- (3) $\mathbf{0}$ is in H .

Activity 22.3. Is the given subset H a subspace of the indicated vector space V ? Verify your answer.

- (a) V is any vector space and $H = \{\mathbf{0}\}$
- (b) $V = M_{2 \times 2}$, the vector space of 2×2 matrices and

$$H = \left\{ \begin{bmatrix} 2x & y \\ 0 & x \end{bmatrix} \mid x \text{ and } y \text{ are scalars} \right\}.$$
- (c) $V = \mathbb{P}_2$, the vector space of all polynomials of degree less than or equal to 2 and $H = \{2at^2 + 1 \mid a \text{ is a scalar}\}$.
- (d) $V = \mathbb{P}_2$ and $H = \{at \mid a \text{ is a scalar}\} \cup \{bt^2 \mid b \text{ is a scalar}\}$.
- (e) $V = \mathcal{F}$ and $H = \mathbb{P}_2$.

There is an interesting subspace relationship between the spaces $\mathbb{P}_1, \mathbb{P}_2, \mathbb{P}_3, \dots$ and \mathbb{P} . For every i , \mathbb{P}_i is a subspace of \mathbb{P} . Furthermore, \mathbb{P}_1 is a subspace of \mathbb{P}_2 , \mathbb{P}_2 is a subspace of \mathbb{P}_3 , and so on. Note however that a similar relationship does NOT hold for \mathbb{R}^n , even though \mathbb{P}_i looks like \mathbb{R}^{i+1} . For example, \mathbb{R}^1 is NOT a subspace of \mathbb{R}^2 . Similarly, \mathbb{R}^2 is NOT a subspace of \mathbb{R}^3 . Since the vectors in different \mathbb{R}^n 's are of different sizes, none of the \mathbb{R}^i 's is a subset of another \mathbb{R}^n with $i \neq n$, and hence, \mathbb{R}^i is not a subspace of \mathbb{R}^n when $i < n$.

The Subspace Spanned by a Set of Vectors

In \mathbb{R}^n we showed that the span of any set of vectors forms a subspace of \mathbb{R}^n . The same is true in any vector space. Recall that the span of a set of vectors in \mathbb{R}^n is the set of all linear combinations of those vectors. So before we can discuss the span of a set of vectors in a vector space, we need to extend the definition of linear combinations to vector spaces (compare to Definitions 4.4 and 4.6).



Definition 22.5. Let V be a vector space. A **linear combination** of vectors $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k$ in V is a vector of the form

$$x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + \cdots + x_k\mathbf{v}_k,$$

where x_1, x_2, \dots, x_k are scalars. The **span** of the vectors $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k$ is the collection of all linear combinations of $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k$. That is,

$$\text{Span}\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k\} = \{x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + \cdots + x_k\mathbf{v}_k \mid x_1, x_2, \dots, x_k \text{ are scalars}\}.$$

The argument that the span of any finite set of vectors in a vector space forms a subspace is the same as we gave for the span of a set of vectors in \mathbb{R}^n (see Theorem 12.5). The proof is left for the exercises.

Theorem 22.6. Given a vector space V and vectors $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_m$ in V , $\text{Span}\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_m\}$ is a subspace of V .

The subspace $\text{Span}\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_m\}$ is called the *subspace of V spanned by $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_m$* .

Activity 22.4.

- (a) Let $H = \{a_2t^2 - a_1t : a_2 \text{ and } a_1 \text{ are real numbers}\}$. Note that H is a subset of \mathbb{P}_2 . Find two vectors $\mathbf{v}_1, \mathbf{v}_2$ in \mathbb{P}_2 so that $H = \text{Span}\{\mathbf{v}_1, \mathbf{v}_2\}$ and hence conclude that H is a subspace of \mathbb{P}_2 . (Note that the vectors $\mathbf{v}_1, \mathbf{v}_2$ are not unique.)
- (b) Let $p_1(t) = 1 - t^2$ and $p_2(t) = 1 + t^2$, and let $S = \{p_1(t), p_2(t)\}$ in \mathbb{P}_2 . Is the polynomial $q(t) = 3 - 2t^2$ in $\text{Span } S$? (Hint: Create a matrix equation of the form $A\mathbf{x} = \mathbf{b}$ by setting up an appropriate polynomial equation involving $p_1(t), p_2(t)$ and $q(t)$. Under what conditions on A is the system $A\mathbf{x} = \mathbf{b}$ consistent?)
- (c) With S as in part (b), describe as best you can the subspace $\text{Span } S$ of \mathbb{P}_2 .

Given a subspace H , the set S such that $H = \text{Span } S$ is called a *spanning set* of H . In order to determine if a set $S = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k\}$ is a spanning set for H , all we need to do is to show that for every \mathbf{b} in H , the equation

$$x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + \cdots + x_k\mathbf{v}_k = \mathbf{b}$$

has a solution. We will see important uses of special spanning sets called bases in the rest of this chapter.

Examples

What follows are worked examples that use the concepts from this section.

Example 22.7. Determine if each of the following sets is a vector space.

- (a) $V = \{(x, y, z) : x, y, z \in \mathbb{R}\}$ with addition and multiplication by scalars defined by

$$(a, b, c) \oplus (x, y, z) = (a + x, c + z, b + y) \text{ and } k(x, y, z) = (kx, kz, ky),$$

where (a, b, c) and (x, y, z) are in V and $k \in \mathbb{R}$



- (b) $V = \{x \in \mathbb{R} : x > 0\}$ with addition \oplus and multiplication by scalars defined by

$$x \oplus y = xy \text{ and } kx = x^k,$$

where x and y are in V , $k \in \mathbb{R}$, and xy is the standard product of x and y

- (c) The set W of all 2×2 matrices of the form $\begin{bmatrix} a & 0 \\ b & 0 \end{bmatrix}$ where a and b are real numbers using the standard addition and multiplication by scalars on matrices.
- (d) The set W of all functions f from \mathbb{R} to \mathbb{R} such that $f(0) \geq 0$ using the standard addition and multiplication by scalars on functions.

Example Solution.

- (a) We consider the vector space properties in Definition 22.1. Let (a, b, c) , (u, v, w) , and (x, y, z) be in V and let $k, m \in \mathbb{R}$. By the definition of addition and multiplication by scalars, both $(a, b, c) + (x, y, z)$ and $k(x, y, z)$ are in V . Note also that

$$\begin{aligned} (a, b, c) \oplus (x, y, z) &= (a+x, c+z, b+y) \\ &= (x+a, z+c, y+b) \\ &= (x, y, z) \oplus (a, b, c), \end{aligned}$$

and so addition is commutative in V .

Since

$$((1, 1, 0) \oplus (0, 1, 1)) \oplus (0, 0, 1) = (1, 1, 2) \oplus (0, 0, 1) = (1, 3, 1)$$

and

$$(1, 1, 0) \oplus ((0, 1, 1) \oplus (0, 0, 1)) = (1, 1, 0) \oplus (0, 2, 1) = (1, 1, 3),$$

we see that addition is not associative and conclude that V is not a vector space. At this point we can stop since we have shown that V is not a vector space.

- (b) We consider the vector space properties in Definition 22.1. Let x, y , and z be in V and let $k, m \in \mathbb{R}$. Since x and y are both positive real numbers, we know that xy is a positive real number. Thus, $x \oplus y \in V$ and V is closed under its addition. Also, x^k is a positive real number, so $x^k \in V$ as well.

Now

$$x \oplus y = xy = yx = y \oplus x$$

and addition is commutative in V .

Also,

$$(x \oplus y) \oplus z = (xy) \oplus z = (xy)z = x(yz) = x \oplus (yz) = x \oplus (y \oplus z)$$

and addition is associative in V .

Since

$$1 \oplus x = 1x = x,$$



V contains an additive identity, which is 1. The fact that x is a positive real number implies that $\frac{1}{x}$ is a positive real number. Thus, $\frac{1}{x} \in V$ and

$$x \oplus \frac{1}{x} = x \left(\frac{1}{x} \right) = 1$$

and V contains an additive inverse for each of its elements.

We have that

$$\begin{aligned} (k+m)x &= x^{k+m} = x^k x^m = x^k \oplus x^m = k(x) \oplus m(x), \\ k(x \oplus y) &= k(xy) = x^k y^k = x^k \oplus y^k = k(x) \oplus k(y), \\ (km)x &= x^{km} = (x^m)^k = (m(x))^k = k(m(x)) \\ 1x &= x^1 = x. \end{aligned}$$

So V satisfies all of the properties of a vector space.

- (c) Recall that $\mathcal{M}_{2 \times 2}$ is a vector space using the standard addition and multiplication by scalars on matrices. Any matrix of the form $\begin{bmatrix} a & 0 \\ b & 0 \end{bmatrix}$ can be written as

$$\begin{bmatrix} a & 0 \\ b & 0 \end{bmatrix} = a \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + b \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}.$$

So $W = \text{Span} \left\{ \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \right\}$ and W is a subspace of $\mathcal{M}_{2 \times 2}$. Thus, W is a vector space.

- (d) We will show that W is not a vector space. Let $f : \mathbb{R} \rightarrow \mathbb{R}$ be defined by $f(x) = 1$. Then $f(0) \geq 0$ and $f \in W$. However, if $h = (-1)f$, then $h(0) = (-1)f(0) = -1$ and $h \notin W$. It follows that W is not closed under multiplication by scalars and W is not a vector space.

Example 22.8. Let V be a vector space and \mathbf{u} and \mathbf{v} vectors in V . Also, let a and b be scalars. You may use the result of Exercise 4 that $c\mathbf{0} = \mathbf{0}$ for any scalar c in any vector space.

- (a) If $a\mathbf{v} = b\mathbf{v}$ and $\mathbf{v} \neq \mathbf{0}$, must $a = b$? Use the properties of a vector space or provide a counterexample to justify your answer.
- (b) If $a\mathbf{u} = b\mathbf{v}$ and $a \neq 0$, must $\mathbf{u} = \mathbf{v}$? Use the properties of a vector space or provide a counterexample to justify your answer.
- (c) If $a\mathbf{u} = b\mathbf{v}$, must $a = b$ and $\mathbf{u} = \mathbf{v}$? Use the properties of a vector space or provide a counterexample to justify your answer.

Example Solution.

- (a) We will show that this statement is true. Suppose $a\mathbf{v} = b\mathbf{v}$ and $\mathbf{v} \neq \mathbf{0}$. Then $\mathbf{0} = a\mathbf{v} - b\mathbf{v} = (a - b)\mathbf{v}$. If $a = b$, then we are done. So suppose $a \neq b$. Then $a - b \neq 0$ and $\frac{1}{a-b}$ is a real



number. Then

$$\begin{aligned}\frac{1}{a-b}\mathbf{0} &= \frac{1}{a-b}((a-b)\mathbf{v}) \\ \mathbf{0} &= \left(\frac{1}{a-b}(a-b)\right)\mathbf{v} \\ \mathbf{0} &= \mathbf{v}.\end{aligned}$$

But we assumed that $\mathbf{v} \neq \mathbf{0}$, so we can conclude that $a = b$ as desired.

- (b) We will show that this statement is true. Suppose $a\mathbf{u} = a\mathbf{v}$ and $a \neq 0$. Then $\mathbf{0} = a\mathbf{u} - a\mathbf{v} = a(\mathbf{u} - \mathbf{v})$. Since $a \neq 0$, we know that $\frac{1}{a}$ is a real number. Thus,

$$\begin{aligned}\frac{1}{a}\mathbf{0} &= \frac{1}{a}(a(\mathbf{u} - \mathbf{v})) \\ \mathbf{0} &= \left(\frac{1}{a}a\right)(\mathbf{u} - \mathbf{v}) \\ \mathbf{0} &= \mathbf{u} - \mathbf{v} \\ \mathbf{u} &= \mathbf{v}.\end{aligned}$$

- (c) We will demonstrate that this statement is false with a counterexample. Let $a = 1$, $b = 2$, $\mathbf{u} = [2 \ 0]^\top$ and $\mathbf{v} = [1 \ 0]^\top$ in \mathbb{R}^2 . Then

$$a\mathbf{u} = 1[2 \ 0]^\top = [2 \ 0]^\top = 2[1 \ 0]^\top = b\mathbf{v},$$

but $a \neq b$ and $\mathbf{u} \neq \mathbf{v}$.

Summary

- A set V on which an operation of addition and a multiplication by scalars is defined is a vector space if for all \mathbf{u} , \mathbf{v} , and \mathbf{w} in V and all scalars a and b :
 - (1) $\mathbf{u} + \mathbf{v}$ is an element of V ,
 - (2) $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$,
 - (3) $(\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w})$,
 - (4) there is a zero vector $\mathbf{0}$ in V so that $\mathbf{u} + \mathbf{0} = \mathbf{u}$,
 - (5) for each \mathbf{x} in V there is an element \mathbf{y} in V so that $\mathbf{x} + \mathbf{y} = \mathbf{0}$,
 - (6) $a\mathbf{u}$ is an element of V ,
 - (7) $(a + b)\mathbf{u} = a\mathbf{u} + b\mathbf{u}$,
 - (8) $a(\mathbf{u} + \mathbf{v}) = a\mathbf{u} + a\mathbf{v}$,
 - (9) $(ab)\mathbf{u} = a(b\mathbf{u})$,
 - (10) $1\mathbf{u} = \mathbf{u}$.
- A subset H of a vector space V is a subspace of V if



- (1) whenever \mathbf{u} and \mathbf{v} are in H it is also true that $\mathbf{u} + \mathbf{v}$ is in H ,
- (2) whenever \mathbf{u} is in H and a is a scalar it is also true that $a\mathbf{u}$ is in H ,
- (3) $\mathbf{0}$ is in H .
- A linear combination of vectors $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k$ in a vector space V is a vector of the form

$$x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + \cdots + x_k\mathbf{v}_k,$$

where x_1, x_2, \dots, x_k are scalars.

- The span of the vectors $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k$ in a vector space V is the collection of all linear combinations of $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k$. That is,

$$\text{Span}\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k\} = \{x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + \cdots + x_k\mathbf{v}_k : x_1, x_2, \dots, x_k \text{ are scalars}\}.$$

- The span of any finite set of vectors in a vector space V is always a subspace of V .
- This concept of vector space is important because there are many different types of sets (e.g., $\mathbb{R}^n, \mathcal{M}_{m \times n}, \mathbb{P}_n, \mathcal{F}$) that have similar structure, and we can relate them all as members of this larger collection of vector spaces.

Exercises

- (1) The definition of a vector space only states the existence of a zero vector and does not say how many zero vectors the space might have. In this exercise we show that the zero vector in a vector space is unique. To show that the zero vector is unique, we assume that two vectors $\mathbf{0}_1$ and $\mathbf{0}_2$ have the zero vector property.
 - (a) Using the fact that $\mathbf{0}_1$ is a zero vector, what vector is $\mathbf{0}_1 + \mathbf{0}_2$?
 - (b) Using the fact that $\mathbf{0}_2$ is a zero vector, what vector is $\mathbf{0}_1 + \mathbf{0}_2$?
 - (c) How do we conclude that the zero vector is unique?
- (2) The definition of a vector space only states the existence of an additive inverse for each vector in the space, but does not say how many additive inverses a vector can have. In this exercise we show that the additive inverse of a vector in a vector space is unique. To show that a vector \mathbf{v} has only one additive inverse, we suppose that \mathbf{v} has two additive inverses, \mathbf{u} and \mathbf{w} , and demonstrate that $\mathbf{u} = \mathbf{w}$.
 - (a) What equations must \mathbf{u} and \mathbf{w} satisfy if \mathbf{u} and \mathbf{w} are additive inverses of \mathbf{v} ?
 - (b) Use the equations from part (a) to show that $\mathbf{u} = \mathbf{w}$. Clearly identify all vector space properties you use in your argument.
- (3) Let V be a vector space and \mathbf{v} a vector in V . In all of the vector spaces we have seen to date, the additive inverse of the vector \mathbf{v} is equal to the scalar multiple $(-1)\mathbf{v}$. This seems reasonable, but it is important to note that this result is not stated in the definition of a vector



space, so this it is something that we need to verify. To show that $(-1)\mathbf{v}$ is an additive inverse of the vector \mathbf{v} , we need to demonstrate that

$$\mathbf{v} + (-1)\mathbf{v} = \mathbf{0}.$$

Verify this equation, explicitly stating which properties you use at each step.

- (4) It is reasonable to expect that if c is any scalar and $\mathbf{0}$ is the zero vector in a vector space V , then $c\mathbf{0} = \mathbf{0}$. Use the fact that $\mathbf{0} + \mathbf{0} = \mathbf{0}$ to prove this statement.
- (5) Let W_1, W_2 be two subspaces of a vector space V . Determine whether $W_1 \cap W_2$ and $W_1 \cup W_2$ are subspaces of V . Justify each answer clearly.
- (6) Find three vectors $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ to express

$$W = \left\{ \begin{bmatrix} a + 2b + c \\ b - 3c \\ a - c \end{bmatrix} : a, b, c \text{ in } \mathbb{R} \right\}$$

as $\text{Span}\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$. How does this justify why W is a subspace of \mathbb{R}^3 ?

- (7) Find three vectors $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ to express

$$W = \left\{ \begin{bmatrix} a + b & a - 2c \\ 3b + c & a + b - c \end{bmatrix} : a, b, c \text{ in } \mathbb{R} \right\}$$

as $\text{Span}\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$. How does this justify why W is a subspace of $\mathcal{M}_{2 \times 2}$?

- (8) Let \mathcal{F} be the set of all functions from \mathbb{R} to \mathbb{R} , where we define the sum of two functions f and g in \mathcal{F} as the function $f + g$ satisfying

$$(f + g)(x) = f(x) + g(x)$$

for all real numbers x , and the scalar multiple cf of the function f by the scalar c to be the function satisfying

$$(cf)(x) = cf(x)$$

for all real numbers x . Show that \mathcal{F} is a vector space using these operations.

- (9) Prove Theorem 22.6. (Hint: Compare to Theorem 12.5).
- (10) Determine if each of the following sets of elements is a vector space or not. If appropriate, you can identify a set as a subspace of another vector space, or as a span of a collection of vectors to shorten your solution.
 - (a) A line through the origin in \mathbb{R}^n .
 - (b) The first quadrant in \mathbb{R}^2 .
 - (c) The set of vectors $\left\{ \begin{bmatrix} a \\ 0 \end{bmatrix} : a \text{ in } \mathbb{Z} \right\}$.
 - (d) The set of all differentiable functions from \mathbb{R} to \mathbb{R} .



- (e) The set of all functions from \mathbb{R} to \mathbb{R} which are increasing for every x . (Assume that a function f is increasing if $f(a) > f(b)$ whenever $a > b$.)
- (f) The set of all functions f from \mathbb{R} to \mathbb{R} for which $f(c) = 0$ for some fixed c in \mathbb{R} .
- (g) The set of polynomials of the form $a + bt$, where $a + b = 0$.
- (h) The set of all upper triangular 4×4 real matrices.
- (i) The set of complex numbers \mathbb{C} where scalar multiplication is defined as multiplication by real numbers.
- (11) A reasonable way to extend the idea of the vector space \mathbb{R}^n to infinity is to let \mathbb{R}^∞ be the set of all sequences of real numbers. Define addition and multiplication by scalars on \mathbb{R}^∞ by

$$\{x_n\} + \{y_n\} = \{x_n + y_n\} \text{ and } c\{x_n\} = \{cx_n\}$$

where $\{x_n\}$ denotes the sequence x_1, x_2, x_3, \dots , $\{y_n\}$ denotes the sequence y_1, y_2, y_3, \dots and c is a scalar.

- (a) Show that \mathbb{R}^∞ is a vector space using these operations.
- (b) Is the set of sequences that have infinitely many zeros a subspace of \mathbb{R}^∞ ? Verify your answer.
- (c) Is the set of sequences which are eventually zero a subspace of \mathbb{R}^∞ ? Verify your answer. (A sequence $\{x_n\}$ is *eventually zero* if there is an index k_0 such that $x_n = 0$ whenever $n \geq k_0$.)
- (d) Is the set of decreasing sequences a subspace of \mathbb{R}^∞ ? Verify your answer. (A sequence $\{x_n\}$ is *decreasing* if $x_{n+1} \leq x_n$ for each n .)
- (e) Is the set of sequences in \mathbb{R}^∞ that have limits at infinity a subspace of \mathbb{R}^∞ ?
- (f) Let ℓ^2 be the set of all square summable sequences in \mathbb{R}^∞ , that is sequences $\{x_n\}$ so that $\sum_{k=1}^{\infty} x_k^2$ is finite. So, for example, the sequence $\{\frac{1}{n}\}$ is in ℓ^2 . Show that ℓ^2 is a subspace of \mathbb{R}^∞ (the set ℓ^2 is an example of what is called a *Hilbert space* by defining the inner product $\langle \{x_n\}, \{y_n\} \rangle = \sum_{n=1}^{\infty} x_n y_n$). (Hint: show that $2u^2 + 2v^2 - (u + v)^2 \geq 0$ for any real numbers u and v .)

- (12) Given two subspaces H_1, H_2 of a vector space V , define

$$H_1 + H_2 = \{\mathbf{w} \mid \mathbf{w} = \mathbf{u} + \mathbf{v} \text{ where } \mathbf{u} \text{ in } H_1, \mathbf{v} \text{ in } H_2\}.$$

Show that $H_1 + H_2$ is a subspace of V containing both H_1, H_2 as subspaces. The space $H_1 + H_2$ is the sum of the subspaces H_1 and H_2 .

- (13) Label each of the following statements as True or False. Provide justification for your response.
- (a) **True/False** The intersection of any two subspaces of V is also a subspace.
- (b) **True/False** The union of any two subspaces of V is also a subspace.
- (c) **True/False** If H is a subspace of a vector space V , then $-H = \{(-1)\mathbf{v} : \mathbf{v} \text{ in } H\}$ is equal to H .



- (d) **True/False** If \mathbf{v} is a nonzero vector in H , a subspace of \mathbb{R}^n , then H contains the line through the origin and \mathbf{v} in \mathbb{R}^n .
- (e) **True/False** If $\mathbf{v}_1, \mathbf{v}_2$ are nonzero, non-parallel vectors in H , a subspace of \mathbb{R}^n , then H contains the plane through the origin, \mathbf{v}_1 and \mathbf{v}_2 in \mathbb{R}^n .
- (f) **True/False** The smallest subspace in \mathbb{R}^n containing a vector \mathbf{v} is a line through the origin.
- (g) **True/False** The largest subspace of V is V .
- (h) **True/False** The space \mathbb{P}_1 is a subspace of \mathbb{P}_n for $n \geq 1$.
- (i) **True/False** The set of constant functions from \mathbb{R} to \mathbb{R} is a subspace of \mathcal{F} .
- (j) **True/False** The set of all polynomial functions with rational coefficients is a subspace of \mathcal{F} .

Project: Hamming Codes and the Hat Puzzle

Recall the hat problem from the beginning of this section. Three players are assigned either a red or blue hat and can only see the colors of the hats of the other players. The goal is to devise a high probability strategy for one player to correctly guess the color of their hat. The players have a 50% chance of winning if one player guesses randomly and all of the others pass. However, the group can do better than 50% with a reasonably simple strategy. There are 2 possibilities for each hat color for a total of $2^3 = 8$ possible distributions of hat colors. Of these, only red-red-red and blue-blue-blue contain only one hat color, so $6/8$ of $3/4$ of the possible hat distributions have two hats of one color and one of the other color. So if a player sees two hats of the same color, that player guesses the other color and passes otherwise. This gives a 75% chance of winning. This strategy will only work for three players, though. We want to develop an effective strategy that works for larger groups of players.

There is a strategy, based on *Hamming codes* that can be utilized when the number of players is of the form $2^k - 1$ with $k \geq 2$. This strategy will provide a winning probability of

$$1 - 2^{-k}.$$

Note that as $k \rightarrow \infty$, this probability has a limit of 1. Note also that if $k = 2$ (so that there are 3 players), then the probability is $\frac{3}{4}$ or 75% – the same strategy we came up with earlier.

To understand this strategy, we need to build a slightly different kind of vector space than we have seen until now, one that is based on a binary choice of red or blue. To do so, we identify the hat colors with numbers – 0 for red and 1 for blue. So let $\mathbb{F} = \{0, 1\}$. Assume there are $n = 2^k - 1$ players for some integer $k \geq 2$. We can then view a distribution of hats among the $n = 2^k - 1$ players as a vector with n components from \mathbb{F} . That is,

$$\mathbb{F}^n = \{[\alpha_1 \ \alpha_2 \ \cdots \ \alpha_n]^\top : \alpha_i \in \mathbb{F}\}.$$

We can give some structure to both \mathbb{F} and \mathbb{F}^n by noting that we can define addition and multi-



plication in \mathbb{F} by

$$\begin{aligned} 0 + 0 &= 0, & 0 + 1 = 1 + 0 &= 1, & 1 + 1 &= 0 \\ 0 \cdot 0 &= 0, & 0 \cdot 1 = 1 \cdot 0 &= 0, & 1 \cdot 1 &= 1. \end{aligned}$$

Project Activity 22.1. Show that \mathbb{F} has the same structure as \mathbb{R} . That is, show that for all x, y , and z in \mathbb{F} , the following properties are satisfied.

- (a) $x + y \in \mathbb{F}$ and $xy \in \mathbb{F}$
- (b) $x + y = y + x$ and $xy = yx$
- (c) $(x + y) + z = x + (y + z)$ and $(xy)z = x(yz)$
- (d) There is an element 0 in \mathbb{F} such that $x + 0 = x$
- (e) There is an element 1 in \mathbb{F} such that $(1)x = x$
- (f) There is an element $-x$ in \mathbb{F} such that $x + (-x) = 0$
- (g) If $x \neq 0$, there is an element $\frac{1}{x}$ in \mathbb{F} such that $x(\frac{1}{x}) = 1$
- (h) $x(y + z) = (xy) + (xz)$

Project Activity 22.1 shows that \mathbb{F} has the same properties as \mathbb{R} – that is that \mathbb{F} is a field. Until now, we have worked with vector spaces whose scalars come from the set of real numbers, but that is not necessary. None of the results we have discovered so far about vector spaces require our scalars to come from \mathbb{R} . In fact, we can replace \mathbb{R} with any field and all of the same vector space properties hold. It follows that $V = \mathbb{F}^n$ is a vector space over \mathbb{F} . As we did in \mathbb{R}^n , we define the standard unit vectors $\mathbf{e}_1 = [1 \ 0 \ 0 \ \cdots \ 0]^\top$, $\mathbf{e}_2 = [0 \ 1 \ 0 \ 0 \ \cdots \ 0]^\top$, ..., $\mathbf{e}_n = [0 \ 0 \ 0 \ \cdots \ 0 \ 1]^\top$ in $V = \mathbb{F}^n$.

Now we return to the hat puzzle. We have $n = 2^k - 1$ players. Label the players 1, 2, ..., n . We can now represent a random placements of hats on heads as a vector $\mathbf{v} = [\alpha_1 \ \alpha_2 \ \cdots \ \alpha_n]^\top$ in $V = \mathbb{F}^n$, where $\alpha_i = 0$ in the i th entry represents a red hat and $\alpha_i = 1$ a blue hat on player i . Since player i can see all of the other hats, from player i 's perspective the distribution of hats has the form

$$\mathbf{v} = \mathbf{v}_i + \beta_i \mathbf{e}_i,$$

where β_i is the unknown color of hat on player i 's head and

$$\mathbf{v}_i = [\alpha_1 \ \alpha_2 \ \cdots \ \alpha_{i-1} \ 0 \ \alpha_{i+1} \ \cdots \ \alpha_n]^\top.$$

In order to analyze the vectors \mathbf{v} from player i 's perspective and to devise an effective strategy, we will partition the set V into an appropriate disjoint union of subsets.

To provide a different way to look at players, we will use a subspace of V . Let W be a subspace of V that has a basis of k vectors. The elements of W are the linear combinations of k basis vectors, and each basis vector in a linear combination has 2 possibilities for its weight (from \mathbb{F}). Thus, W contains exactly $2^k = n + 1$ vectors. We can then use the $n = 2^k - 1$ nonzero vectors in W to represent our players. Each distribution of hats can be seen as a linear combination of the vectors



in W . Let $\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_{2^k-1}$ be the nonzero vectors in W . We then define a function $\varphi : V \rightarrow W$ as

$$\varphi([\alpha_1 \ \alpha_2 \ \cdots \ \alpha_n]^\top) = \sum_{i=1}^n \alpha_i \mathbf{w}_i$$

that identifies a distribution of hats with a vector in W . The subspace that we need to devise our strategy is what is called a Hamming code.

Project Activity 22.2. Let

$$H = \left\{ [\alpha_1 \ \alpha_2 \ \cdots \ \alpha_n]^\top \in V : \sum_{i=1}^n \alpha_i \mathbf{w}_i = \mathbf{0} \right\}.$$

Show that H is a subspace of V . (The subspace H is called the $(2^k - 1, 2^k - k - 1)$ Hamming code (where the first component is the number of elements in a basis for V and the second the number of elements in a basis for H). Hamming codes are examples of linear codes – those codes that are subspaces of the larger vector space.)

Now for each i between 0 and n we define $H_i = \mathbf{e}_i + H$ as

$$H_i = \mathbf{e}_i + H = \{\mathbf{e}_i + \mathbf{h} : \mathbf{h} \in H\},$$

where we let $\mathbf{e}_0 = \mathbf{0}$. The sets H_i are called *cosets* of H .

Project Activity 22.3. To complete our strategy for the hat puzzle, we need to know some additional information about the sets H_i .

- (a) Show that the sets H_i are disjoint. That is, show that $H_i \cap H_j = \emptyset$ if $i \neq j$. (Hint: If $\mathbf{v} \in H_i$ and $\mathbf{v} \in H_j$, what can we say about $\mathbf{e}_i - \mathbf{e}_j$?)
- (b) Since $H_i \subseteq V$ for each i , it follows that $\bigcup_{i=0}^n H_i \subseteq V$. Now we show that $V = \bigcup_{i=0}^n H_i$ by demonstrating that $\bigcup_{i=0}^n H_i$ has exactly the same number of elements as V . We will need one fact for our argument. We will see in a later section that H has a basis of $n - k$ elements, so the number of elements in H is 2^{n-k} .
 - i. Since the sets H_i are disjoint, the number of elements in $\bigcup_{i=1}^n H_i$ is equal to the sum of the number of elements in each H_i . Show that each H_i has the same number of elements as H .
 - ii. Now use the fact that the number of elements in $\bigcup_{i=0}^n H_i$ is equal to the sum of the number of elements in each H_i to argue that $V = \bigcup_{i=0}^n H_i$.

The useful idea from Project Activity 22.3 is that any hat distribution in V is in exactly one of the sets H_i . Recall that a hat distribution $\mathbf{v} = [\alpha_1 \ \alpha_2 \ \cdots \ \alpha_n]^\top$ in V can be written from player i 's perspective as

$$\mathbf{v} = \mathbf{v}_i + \beta_i \mathbf{e}_i,$$

where $\mathbf{v}_i = [\alpha_1 \ \alpha_2 \ \cdots \ \alpha_{i-1} \ 0 \ \alpha_{i+1} \ \cdots \ \alpha_n]^\top$. Our strategy for the hat game can now be revealed.

- If $\mathbf{v}_i + \beta_i \mathbf{e}_i$ is not in H for either choice of β_i , then player i should pass.



- If $\mathbf{v}_i + \beta_i \mathbf{e}_i$ is in H , then player i guesses $1 + \beta_i$.

Project Activity 22.4. Let us analyze this strategy.

- Explain why every player guesses wrong if \mathbf{v} is in H .
- Now we see determine that our strategy is a winning strategy for all hat distributions \mathbf{v} that are not in H . First we need to know that these two options are the only ones. That is, show that it is not possible for $\mathbf{v}_i + \beta_i \mathbf{e}_i$ to be in H for both choices of β_i .
- Now we want to demonstrate that this is a winning strategy if $\mathbf{v} \notin H$. That is, at least one player guesses a correct hat color and no one else guesses incorrectly. So assume $\mathbf{v} \notin H$.
 - We know that $\mathbf{v} \in H_i$ for some unique choice of i , so let $\mathbf{v} = \mathbf{e}_i + h$ for some $h \in H$. Explain why player i can correctly choose color $1 + \alpha_i$.
 - Finally, we need to argue that every player except player i must pass. So consider player j , with $j \neq i$. Recall that

$$\mathbf{v} = \mathbf{v}_j + \alpha_j \mathbf{e}_j.$$

Analyze our strategy and the conditions under which player j does not pass. Show that this leads to a contradiction.

Project Activity 22.4 completes our analysis of this strategy and shows that our strategy results in a win with probability

$$1 - \frac{|H|}{|V|} = 1 - \frac{2^{2^k-k-1}}{2^{2^k-1}} = 1 - 2^{-k}.$$

Section 23

Bases for Vector Spaces

Focus Questions

By the end of this section, you should be able to give precise and thorough answers to the questions listed below. You may want to keep these questions in mind to focus your thoughts as you complete the section.

- What does it mean for a set $\{v_1, v_2, \dots, v_k\}$ of vectors in a vector space V to be linearly independent?
- What is another equivalent characterization of a linearly independent set?
- What does it mean for a set $\{v_1, v_2, \dots, v_k\}$ of vectors in a vector space V to be linearly dependent?
- Describe another characterization of a linearly dependent set.
- What is a basis for a vector space V ?
- What makes a basis for a vector space useful?
- How can we find a basis for a vector space V ?

Application: Image Compression

If you painted a picture with a sky, clouds, trees, and flowers, you would use a different size brush depending on the size of the features. Wavelets are like those brushes.

– Ingrid Daubechies

The advent of the digital age has presented many new opportunities for the collection, analysis, and dissemination of information. Along with these opportunities come new difficulties as well. All of this digital information must be stored in some way and be retrievable in an efficient manner. One collection of tools that is used to deal with these problems is wavelets. For example, The FBI fingerprint files contain millions of cards, each of which contains 10 rolled fingerprint impressions. Each card produces about 10 megabytes of data. To store all of these cards would require an enor-

mous amount of space, and transmitting one full card over existing data lines is slow and inefficient. Without some sort of image compression, a sortable and searchable electronic fingerprint database would be next to impossible. To deal with this problem, the FBI adopted standards for fingerprint digitization using a wavelet compression standard.

Another problem with electronics is noise. Noise can be a big problem when collecting and transmitting data. Wavelet decomposition filters data by averaging and detailing. The detailing coefficients indicate where the details are in the original data set. If some details are very small in relation to others, eliminating them may not substantially alter the original data set. Similar ideas may be used to restore damaged audio,¹ video, photographs, and medical information.²

We will consider wavelets as a tool for image compression. The basic idea behind using wavelets to compress images is that we start with a digital image, made up of pixels. Each pixel can be assigned a number or a vector (depending on the makeup of the image). The image can then be represented as a matrix (or a set of matrices) M , where each entry in M represents a pixel in the image. As a simple example, consider the 16×16 image of a flower as shown at left in Figure 23.1. (We will work with small images like this to make the calculations more manageable, but the ideas work for any size image. We could also extend our methods to consider color images, but for the sake of simplicity we focus on grayscale.) This flower image is a gray-scale image, so each

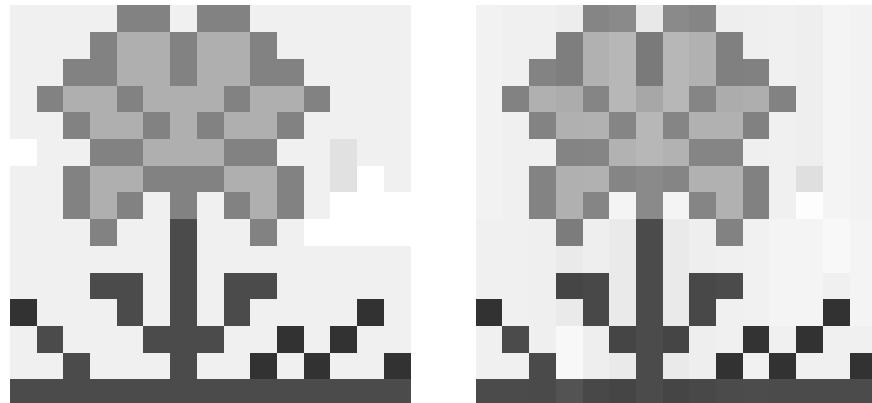


Figure 23.1: Left: A 16 by 16 pixel image. Right: The image compressed.

pixel has a numeric representation between 0 and 255, where 0 is black, 255 is white, and numbers

¹see <https://ccrma.stanford.edu/groups/edison/brahms/brahms.html> for a discussion of the denoising of a Brahms recording

²A review of wavelets in biomedical applications. M. Unser, A. Aldroubi. *Proceedings of the IEEE*, Volume: 84, Issue: 4 , Apr 1996



between 0 and 255 represent shades of gray. The matrix for this flower image is

$$\begin{bmatrix} 240 & 240 & 240 & 240 & 130 & 130 & 240 & 130 & 130 & 240 & 240 & 240 & 240 & 240 & 240 & 240 \\ 240 & 240 & 240 & 130 & 175 & 175 & 130 & 175 & 175 & 130 & 240 & 240 & 240 & 240 & 240 & 240 \\ 240 & 240 & 130 & 130 & 175 & 175 & 130 & 175 & 175 & 130 & 130 & 240 & 240 & 240 & 240 & 240 \\ 240 & 130 & 175 & 175 & 130 & 175 & 175 & 130 & 175 & 175 & 130 & 240 & 240 & 240 & 240 & 240 \\ 240 & 240 & 130 & 175 & 175 & 130 & 175 & 175 & 130 & 175 & 175 & 130 & 240 & 240 & 240 & 240 \\ 255 & 240 & 240 & 130 & 130 & 175 & 175 & 130 & 175 & 175 & 130 & 240 & 240 & 240 & 240 & 240 \\ 240 & 240 & 130 & 175 & 175 & 130 & 175 & 175 & 130 & 175 & 175 & 130 & 240 & 225 & 240 & 240 \\ 240 & 240 & 130 & 175 & 175 & 130 & 175 & 175 & 130 & 175 & 175 & 130 & 240 & 225 & 255 & 240 \\ 240 & 240 & 130 & 175 & 175 & 130 & 175 & 175 & 130 & 175 & 175 & 130 & 240 & 255 & 255 & 255 \\ 240 & 240 & 240 & 130 & 240 & 240 & 130 & 240 & 130 & 175 & 130 & 240 & 255 & 255 & 255 & 255 \\ 240 & 240 & 240 & 240 & 240 & 240 & 75 & 240 & 240 & 130 & 240 & 255 & 255 & 255 & 255 & 255 \\ 240 & 240 & 240 & 240 & 240 & 75 & 240 & 240 & 240 & 240 & 240 & 240 & 240 & 240 & 240 & 240 \\ 240 & 240 & 240 & 240 & 75 & 240 & 75 & 240 & 75 & 240 & 75 & 240 & 240 & 240 & 240 & 240 \\ 50 & 240 & 240 & 240 & 75 & 240 & 75 & 240 & 75 & 240 & 240 & 240 & 240 & 240 & 50 & 240 \\ 240 & 75 & 240 & 240 & 240 & 75 & 75 & 240 & 240 & 240 & 50 & 240 & 50 & 240 & 240 & 50 \\ 240 & 240 & 75 & 240 & 240 & 240 & 75 & 240 & 240 & 50 & 240 & 50 & 240 & 240 & 50 & 240 \\ 75 & 75 & 75 & 75 & 75 & 75 & 75 & 75 & 75 & 75 & 75 & 75 & 75 & 75 & 75 & 75 \\ 75 & 75 & 75 & 75 & 75 & 75 & 75 & 75 & 75 & 75 & 75 & 75 & 75 & 75 & 75 & 75 \end{bmatrix} . \quad (23.1)$$

Now we can apply wavelets to the image and compress it. Essentially, wavelets act by averaging and differencing. The averaging creates smaller versions of the image and the differencing keeps track of how far the smaller version is from a previous copy. The differencing often produces many small (close to 0) entries, and so replacing these entries with 0 doesn't have much effect on the image (this is called *thresholding*). By introducing long strings of zeros into our data, we are able to store a (compressed) copy of the image in a smaller amount of space. For example, using a threshold value of 10 produces the flower image shown at right in Figure 23.1.

The averaging and differencing is done with special vectors (wavelets) that form a basis for a suitable function space. More details of this process can be found at the end of this section.

Introduction

In \mathbb{R}^n we defined a basis for a subspace W of \mathbb{R}^n to be a minimal spanning set for W , or a linearly independent spanning set (see Definition 6.6). So to consider the idea of a basis in a vector space, we will need the notion of linear independence in that context.

Since we can add vectors and multiply vectors by scalars in any vector space, and because we have a zero vector in any vector space, we can define linear independence of a finite set of vectors in any vector space as follows (compare to Definition 6.1).

Definition 23.1. A set $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k\}$ of vectors in a vector space V is **linearly independent** if the vector equation

$$x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + \cdots + x_k\mathbf{v}_k = \mathbf{0}$$

for scalars x_1, x_2, \dots, x_k has only the trivial solution

$$x_1 = x_2 = x_3 = \cdots = x_k = 0.$$

If a set of vectors is not linearly independent, then the set is **linearly dependent**.

Alternatively, we say that the vectors $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k$ are linearly independent (or dependent) if the set $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k\}$ is linearly independent (or dependent).

Preview Activity 23.1.



- (1) We can use the tools we developed to determine if a set of vectors in \mathbb{R}^n is linearly independent to answer the same questions for sets of vectors in other vector spaces. For example, consider the question of whether the set $\{1+t, 1-t\}$ in \mathbb{P}_1 is linearly independent or dependent. To answer this question we need to determine if there is a non-trivial solution to the equation

$$x_1(1+t) + x_2(1-t) = 0. \quad (23.2)$$

Note that equation (23.2) can also be written in the form

$$(x_1 + x_2) + (x_1 - x_2)t = 0.$$

- (a) Recall that two polynomials are equal if all coefficients of like powers are the same. By equating coefficients of like power terms, rewrite equation (23.2) as an equivalent system of two equations in the two unknowns x_1 and x_2 , and solve for x_1, x_2 .
- (b) What does your answer to the previous part tell you about the linear independence or dependence of the set $\{1+t, 1-t\}$ in \mathbb{P}_1 ?
- (c) Recall that in \mathbb{R}^n , a set of two vectors is linearly dependent if and only if one of the vectors in the set is a scalar multiple of the other and linearly independent if neither vector is a scalar multiple of the other. Verify your answer to part (c) from a similar perspective in \mathbb{P}_1 .
- (2) We can use the same type of method as in problem (1) to address the question of whether the set

$$\left\{ \begin{bmatrix} 1 & 3 \\ 1 & 2 \end{bmatrix}, \begin{bmatrix} 1 & -9 \\ 1 & 8 \end{bmatrix}, \begin{bmatrix} 1 & -1 \\ 1 & 4 \end{bmatrix} \right\}$$

is linearly independent or dependent in $\mathcal{M}_{2 \times 2}$. To answer this question we need to determine if there is a non-trivial solution to the equation

$$x_1 \begin{bmatrix} 1 & 3 \\ 1 & 2 \end{bmatrix} + x_2 \begin{bmatrix} 1 & -9 \\ 1 & 8 \end{bmatrix} + x_3 \begin{bmatrix} 1 & -1 \\ 1 & 4 \end{bmatrix} = \mathbf{0} \quad (23.3)$$

for some scalars x_1, x_2 , and x_3 . Note that the linear combination on the left side of equation (23.3) has entries

$$\begin{bmatrix} x_1 + x_2 + x_3 & 3x_1 - 9x_2 - x_3 \\ x_1 + x_2 + x_3 & 2x_1 + 8x_2 + 4x_3 \end{bmatrix}.$$

- (a) Recall that two matrices are equal if all corresponding entries are the same. Equate corresponding entries of the matrices in equation (23.3) to rewrite the equation as an equivalent system of four equations in the three unknowns x_1, x_2 , and x_3 .
- (b) Use appropriate matrix tools and techniques to find all solutions to the system from part (a).
- (c) What does the set of solutions to the system from part (a) tell you about the linear independence or dependence of the set

$$\left\{ \begin{bmatrix} 1 & 3 \\ 1 & 2 \end{bmatrix}, \begin{bmatrix} 1 & -9 \\ 1 & 8 \end{bmatrix}, \begin{bmatrix} 1 & -1 \\ 1 & 4 \end{bmatrix} \right\}?$$

- (d) Recall that in \mathbb{R}^n , a set of vectors is linearly dependent if and only if one of the vectors in the set is a linear combination of the others and linearly independent if no vector in the set is a linear combination of the others. Verify your answer to part (c) from a similar perspective in $\mathcal{M}_{2 \times 2}$.
- (3) We will define a basis for a vector space to be a linearly independent spanning set. Which, if any, of the sets in parts (1) and (2) is a basis for its vector space? Explain.

Linear Independence

The concept of linear independence, which we formally defined in Preview Activity 23.1, provides us with a process to determine if there is redundancy in a spanning set to obtain an efficient spanning set.

The definition tells us that a set $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k\}$ of vectors in a vector space V is linearly dependent if there are scalars x_1, x_2, \dots, x_n , not all of which are 0 so that

$$x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + \cdots + x_k\mathbf{v}_k = \mathbf{0}.$$

As examples, we saw in Preview Activity 23.1 that the set $\{1+t, 1-t\}$ is linearly independent in \mathbb{P}_1 . The set $\{1+t, -1+2t+t^2, 1-8t-3t^2\}$, on the other hand, is linearly dependent in \mathbb{P}_2 since $2(1+t) + 3(-1+2t+t^2) + (1-8t-3t^2) = 0$.

In addition to the definition, there are other ways to characterize linearly independent and dependent sets in vector spaces as the next theorems illustrate. These characterizations are the same as those we saw in \mathbb{R}^n , and the proofs are essentially the same as well. The proof of Theorem 23.2 is similar to that of Theorem 6.2 and is left for the exercises.

Theorem 23.2. *A set $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k\}$ of vectors in a vector space V is linearly dependent if and only if at least one of the vectors in the set can be written as a linear combination of the remaining vectors in the set.*

Theorem 23.2 is equivalent to the following theorem that provides the corresponding result for linearly independent sets.

Theorem 23.3. *A set $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k\}$ of vectors in a vector space V is linearly independent if and only if no vector in the set can be written as a linear combination of the remaining vectors in the set.*

One consequence of Theorems 23.2 and 23.3 is that if a spanning set is linearly dependent, then one of the vectors in the set can be written as a linear combination of the others. In other words, at least one of the vectors is redundant. In that case, we can find a smaller spanning set as the next theorem states. The proof of this theorem is similar to that of Theorem 6.5 and is left for the exercises.

Theorem 23.4. *Let $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k\}$ be a set of vectors in a vector space V . If for some i between 1 and k , \mathbf{v}_i is in $\text{Span}\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_{i-1}, \mathbf{v}_{i+1}, \dots, \mathbf{v}_k\}$, then*

$$\text{Span}\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k\} = \text{Span}\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_{i-1}, \mathbf{v}_{i+1}, \dots, \mathbf{v}_k\}.$$



Bases

A basis for a vector space is a spanning set that is as small as it can be. We already saw how to define bases formally in \mathbb{R}^n . We will now formally define a basis for a vector space and understand why with this definition a basis is a minimal spanning set. Bases are important because any vector in a vector space can be uniquely represented as a linear combination of basis vectors. We will see in later sections that this representation will allow us to identify any vector space with a basis of n vectors with \mathbb{R}^n .

To obtain the formal definition of a basis, which is a minimal spanning set, we consider what additional property makes a spanning set a *minimal* spanning set. As a consequence of Theorem 23.4, if S is a spanning set that is linearly dependent, then we can find a proper subset of S that has the same span. Thus, the set S cannot be a minimal spanning set. However, if S is linearly independent, then no vector in S is a linear combination of the others and we need all of the vectors in S to form the span. This leads us to the following formal characterization of a minimal spanning set, called a *basis*.

Definition 23.5. A **basis** for a vector space V is a subset S of V if

- (1) $\text{Span } S = V$ and
- (2) S is a linearly independent set.

In other words, a basis for a vector space V is a linearly independent spanning set for V . To put it another way, a basis for a vector space is a minimal spanning set for the vector space. Similar reasoning will show that a basis is also a maximal linearly independent set.

The key ideas to take from the previous theorems are:

- A basis for a vector space V is a minimal spanning set for V .
- A basis for V is a subset S of V so that
 - (1) S spans V and
 - (2) S is linearly independent.
- No vector in a basis can be written as a linear combination of the other vectors in the basis.
- If a subset S of a vector space V has the property that one of the vectors in S is a linear combination of the other vectors in S , then S is not a basis for V .

As an example of a basis of a vector space, we saw in Preview Activity 23.1 that the set $S = \{1 - t, 1 + t\}$ is both linearly independent and spans \mathbb{P}_1 , and so S is a basis for \mathbb{P}_1 .

Activity 23.1.

- (a) Is $S = \{1 + t, t, 1 - t\}$ a basis for \mathbb{P}_1 ? Explain.
- (b) Explain why the set $S = \{1, t, t^2, \dots, t^n\}$ is a basis for \mathbb{P}_n . This basis is called the *standard basis* for \mathbb{P}_n .



- (c) Show that the set

$$\left\{ \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \right\}$$

is a basis for $\mathcal{M}_{2 \times 2}$.

It should be noted that not every vector space has a finite basis. For example, the space \mathbb{P} of all polynomials with real coefficients (of any degree) is a vector space, but no finite set of vectors will span \mathbb{P} . In fact, the infinite set $\{1, t, t^2, \dots\}$ is both linearly independent and spans \mathbb{P} , so \mathbb{P} has an infinite basis.

Finding a Basis for a Vector Space

We already know how to find bases for certain vector spaces, namely $\text{Nul } A$ and $\text{Col } A$, where A is any matrix. Finding a basis for a different kind of vector space will require other methods. Since a basis for a vector space is a minimal spanning set, to find a basis for a given vector space we might begin from scratch, starting with a given vector in the space and adding one vector at a time until we have a spanning set.

Activity 23.2. Let $W = \{a + bt + ct^3 \mid a, b, c \text{ are scalars}\}$. We will find a basis of W that contains the polynomial $3 + t - 3t^3$.

- (a) Let $\mathcal{S}_1 = \{3 + t - t^3\}$. Find a polynomial $p(t)$ in W that is not in $\text{Span } \mathcal{S}_1$. Explain why this means that the set \mathcal{S}_1 does not span W .
- (b) Let $\mathcal{S}_2 = \{3 + t - t^3, p(t)\}$. Find a polynomial $q(t)$ that is not in $\text{Span } \mathcal{S}_2$. What does this mean about \mathcal{S}_2 being a possible spanning set of W ?
- (c) Let $\mathcal{S}_3 = \{3 + t - t^3, p(t), q(t)\}$. Explain why the set \mathcal{S}_3 is a basis for W .

Alternatively, we might construct a basis from a known spanning set.

Activity 23.3. Let $W = \left\{ \begin{bmatrix} v+z & w+z \\ x & y \end{bmatrix} \mid v, w, x, y, z \text{ are scalars} \right\}$. Assume that W is a subspace of $\mathcal{M}_{2 \times 2}$.

- (a) Find a set S of five 2×2 matrices that spans W (since W is a span of a set of vectors in $\mathcal{M}_{2 \times 2}$, W is a subspace of $\mathcal{M}_{2 \times 2}$). Without doing any computation, can this set S be a basis for W ? Why or why not?
- (b) Find a subset B of S that is a basis for W .

Activities 23.2 and 23.3 give us two ways of finding a basis for a subspace W of a vector space V , assuming W has a basis with finitely many vectors. One way (illustrated in Activity 23.2) is to start by choosing any non-zero vector w_1 in W . Let $\mathcal{S}_1 = \{w_1\}$. If \mathcal{S}_1 spans W , then \mathcal{S}_1 is a basis for W . If not, there is a vector w_2 in W that is not in $\text{Span } \mathcal{S}_1$. Then $\mathcal{S}_2 = \{w_1, w_2\}$ is a linearly independent set. If $\text{Span } \mathcal{S}_2 = W$, then \mathcal{S}_2 is a basis for W and we are done. If not, repeat the



process. We will show later that this process must stop as long as we know that W has a basis with finitely many vectors.

Another way (illustrated in Activity 23.3) to find a basis for W is to start with a spanning set \mathcal{S}_1 of W . If \mathcal{S}_1 is linearly independent, then \mathcal{S}_1 is a basis for W . If \mathcal{S}_1 is linearly dependent, then one vector in \mathcal{S}_1 is a linear combination of the others and we can remove that vector to obtain a new set \mathcal{S}_2 that also spans W . If \mathcal{S}_2 is linearly independent, then \mathcal{S}_2 is a basis for W . If not, we repeat the process as many times as needed until we arrive until at a subset \mathcal{S}_k of \mathcal{S}_1 that is linearly independent and spans W . We summarize these results in the following theorem.

Theorem 23.6. *Let W be a subspace of a finite-dimensional vector space V . Then*

- (1) *any linearly independent subset of W can be extended to a basis of W ,*
- (2) *any subset of W that spans W can be reduced to a basis of W .*

We conclude this section with the result mentioned in the introduction – that every vector in a vector space with basis \mathcal{B} can be written in one and only one way as a linear combination of basis vectors. The proof is similar to that of Theorem 6.4 and is left to the exercises.

Theorem 23.7. *Let $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$ be vectors in a vector space V that make up a basis \mathcal{B} for V . If \mathbf{u} is a vector in V , then \mathbf{u} can be written in one and only one way as a linear combination of vectors $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$ in \mathcal{B} .*

Examples

What follows are worked examples that use the concepts from this section.

Example 23.8. Let $S = \{1, 1+t, 2-t^2, 1+t+t^2, t-t^2\}$.

- (a) Does S span \mathbb{P}_2 ? Explain.
- (b) Explain why S is not a basis for \mathbb{P}_2 .
- (c) Find a subset of S that is a basis for \mathbb{P}_2 . Explain your reasoning.

Example Solution.

- (a) Let $p(t) = a_0 + a_1t + a_2t^2$ be an arbitrary vector in \mathbb{P}_2 . If $p(t)$ is in $\text{Span } S$, then there are weights c_1, c_2, c_3, c_4 , and c_5 such that

$$a_0 + a_1t + a_2t^2 = c_1(1) + c_2(1+t) + c_3(2-t^2) + c_4(1+t+t^2) + c_5(t-t^2).$$

Equating coefficients of like powers gives us the system

$$\begin{aligned} c_1 + c_2 + 2c_3 + c_4 &= a_0 \\ c_2 + c_4 + c_5 &= a_1 \\ -c_3 + c_4 - c_5 &= a_2. \end{aligned}$$



The reduced row echelon form of the coefficient matrix A is

$$\begin{bmatrix} 1 & 0 & 0 & 2 & -3 \\ 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & -1 & 1 \end{bmatrix}.$$

Since there is a pivot in every row of A , the system $A\mathbf{x} = \mathbf{b}$ is always consistent. We conclude that S does span \mathbb{P}_2 .

- (b) The fact that the coefficient matrix A of our system has non-pivot columns means that each vector in \mathbb{P}_2 can be written in more than one way as a linear combination of vectors in S . This means that S is not linearly independent and so cannot be a basis for \mathbb{P}_2 .
- (c) That the first three columns of A are pivot columns implies that the polynomials 1 , $1 + t$, and $2 - t^2$ are linearly independent. Since there is a pivot in every row of A , the three polynomials 1 , $1 + t$, and $2 - t^2$ also span \mathbb{P}_2 . So $\{1, 1 + t, 2 - t^2\}$ is a subset of S that is a basis for \mathbb{P}_2 .

Example 23.9. Let U be the set of all matrices of real numbers of the form $\begin{bmatrix} u & -u - x \\ 0 & x \end{bmatrix}$ and W be the set of all real matrices of the form $\begin{bmatrix} v & 0 \\ w & -v \end{bmatrix}$.

- (a) Find a basis for U and a basis for W .
- (b) Let $U + W = \{A + B : A \text{ is in } U \text{ and } B \text{ is in } W\}$. Show that $U + W$ is a subspace of $\mathcal{M}_{2 \times 2}$ and find a basis for $U + W$.

Example Solution.

- (a) Every matrix in U has the form

$$\begin{bmatrix} u & -u - x \\ 0 & x \end{bmatrix} = u \begin{bmatrix} 1 & -1 \\ 0 & 0 \end{bmatrix} + x \begin{bmatrix} 0 & -1 \\ 0 & 1 \end{bmatrix}.$$

Let $S_U = \left\{ \begin{bmatrix} 1 & -1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & -1 \\ 0 & 1 \end{bmatrix} \right\}$. Then $U = \text{Span } S_U$ and U is a subspace of $\mathcal{M}_{2 \times 2}$.

If

$$c_1 \begin{bmatrix} 1 & -1 \\ 0 & 0 \end{bmatrix} + c_2 \begin{bmatrix} 0 & -1 \\ 0 & 1 \end{bmatrix} = 0,$$

then $c_1 = c_2 = 0$ and S_U is also linearly independent. This makes S_U a basis for U .

Similarly, every matrix in W has the form

$$\begin{bmatrix} v & 0 \\ w & -v \end{bmatrix} = v \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} + w \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}.$$

Let $S_W = \left\{ \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \right\}$. Then $W = \text{Span } S_W$ and W is a subspace of $\mathcal{M}_{2 \times 2}$. If

$$c_1 \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} + c_2 \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} = 0,$$

then $c_1 = c_2 = 0$ and S_W is also linearly independent. This makes S_W a basis for W .



(b) Every matrix in $U + W$ has the form

$$\begin{aligned} \begin{bmatrix} u & -u-x \\ 0 & x \end{bmatrix} + \begin{bmatrix} v & 0 \\ w & -v \end{bmatrix} &= \begin{bmatrix} u+v & -u-x \\ w & x-v \end{bmatrix} \\ &= u \begin{bmatrix} 1 & -1 \\ 0 & 0 \end{bmatrix} + x \begin{bmatrix} 0 & -1 \\ 0 & 1 \end{bmatrix} \\ &\quad + v \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} + w \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}. \end{aligned}$$

Let $S = \left\{ \begin{bmatrix} 1 & -1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & -1 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \right\}$. Then $U + W = \text{Span } S$ and $U + W$ is a subspace of $\mathcal{M}_{2 \times 2}$. If

$$c_1 \begin{bmatrix} 1 & -1 \\ 0 & 0 \end{bmatrix} + c_2 \begin{bmatrix} 0 & -1 \\ 0 & 1 \end{bmatrix} + c_3 \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} + c_4 \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} = 0,$$

then

$$\begin{aligned} c_1 + c_3 &= 0 \\ -c_1 - c_2 &= 0 \\ c_4 &= 0 \\ c_2 - c_3 &= 0. \end{aligned}$$

The reduced row echelon form of $\begin{bmatrix} 1 & 0 & 1 & 0 \\ -1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & -1 & 0 \end{bmatrix}$ is $\begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$. The vectors that correspond to the pivot columns are linearly independent and span $U + W$, so a basis for $U + W$ is

$$\left\{ \begin{bmatrix} 1 & -1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & -1 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \right\}.$$

Summary

The important idea in this section is that of a basis for a vector space. A basis is a minimal spanning set and another equivalent characterization of the “minimal” property is linear independence.

- A set $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k\}$ of vectors in a vector space V is linearly independent if the vector equation

$$x_1 \mathbf{v}_1 + x_2 \mathbf{v}_2 + \cdots + x_k \mathbf{v}_k = \mathbf{0}$$

for scalars x_1, x_2, \dots, x_k has only the trivial solution

$$x_1 = x_2 = x_3 = \cdots = x_k = 0.$$

If a set of vectors is not linearly independent, then the set is linearly dependent.



- A set $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k\}$ of vectors in a vector space V is linearly independent if and only if none of the vectors in the set can be written as a linear combination of the remaining vectors in the set.
- A set $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k\}$ of vectors in a vector space V is linearly dependent if and only if at least one of the vectors in the set can be written as a linear combination of the remaining vectors in the set.
- A basis for a vector space V is a subset S of V if
 - (1) $\text{Span } S = V$ and
 - (2) S is a linearly independent set.
- A basis is important in that it provides us with an efficient way to represent any vector in the vector space – any vector can be written in one and only one way as a linear combination of vectors in a basis.
- To find a basis of a vector space, we can start with a spanning set S and toss out any vector in S that can be written as a linear combination of the remaining vectors in S . We repeat the process with the remaining subset of S until we arrive at a linearly independent spanning set. Alternatively, we can find a spanning set for the space and remove any vector that is a linear combination of the others in the spanning set. We can repeat this process until we wind up with a linearly independent spanning set.

Exercises

- Determine if the given sets are linearly independent or dependent in the indicated vector space. If dependent, write one of the vectors as a linear combination of the others. If independent, determine if the set is a basis for the vector space.
 - (a) $\{[1 \ 4 \ 6]^T, [2 \ -1 \ 3]^T, [0 \ 1 \ 5]^T\}$ in \mathbb{R}^3
 - (b) $\{1 - 2t^2 + t^3, 3 - t + 4t^3, 2 - 3t\}$ in \mathbb{P}_3
 - (c) $\{1 + t, -1 - 5t + 4t^2 + t^3, 1 + t^2 + t^3, t + 2t^3\}$ in \mathbb{P}_3
 - (d) $\left\{ \begin{bmatrix} 1 & 2 & 0 \\ 0 & 1 & 1 \end{bmatrix}, \begin{bmatrix} 1 & -2 & 0 \\ 0 & -1 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 2 & 0 \\ 0 & 1 & -1 \end{bmatrix} \right\}$ in $\mathcal{M}_{3 \times 2}$.
- Let $S = \{1 + t + t^2, t + t^2, 1 + t, 1 + t^2\}$ in \mathbb{P}_2 .
 - (a) Show that the set S spans \mathbb{P}_2 .
 - (b) Show that the set S is linearly dependent.
 - (c) Find a subset of S that is a basis for \mathbb{P}_2 . Be sure to verify that you have a basis.
- Find two different bases for $\mathcal{M}_{2 \times 2}$. Explain how you know that each set is a basis.
- The set $W = \{at + bt^2 \mid a \text{ and } b \text{ are scalars}\}$ is a subspace of \mathbb{P}_3 .
 - (a) Find a set of vectors in \mathbb{P}_3 that spans W .



- (b) Find a basis for W . Be sure to verify that you have a basis.
- (5) Suppose that the set $\{\mathbf{u}, \mathbf{v}, \mathbf{w}\}$ is a basis for a vector space V . Is the set $\{\mathbf{u}+\mathbf{v}, \mathbf{u}+\mathbf{w}, \mathbf{v}+\mathbf{w}\}$ a basis for V ? Verify your result.
- (6) Determine all scalars c so that the set $\{c^2 + t^2, c + 2t, 1 + t^2\}$ is a basis for \mathbb{P}_2 .
- (7) A symmetric matrix is a matrix A so that $A^\top = A$. Is it possible to find a basis for $\mathcal{M}_{2 \times 2}$ consisting entirely of symmetric matrices? If so, exhibit one such basis. If not, explain why not.
- (8) Find a basis of the subspace of $\mathcal{M}_{2 \times 3}$ consisting of all matrices of the form $\begin{bmatrix} a & b & c \\ d & e & f \end{bmatrix}$ where $c = a - 2d$ and $f = b + 3e$.
- (9) Prove Theorem 23.2. (Hint: Compare to Theorem 6.2.)
- (10) Prove Theorem 23.4. (Hint: Compare to Theorem 6.5.)
- (11) Prove Theorem 23.7. (Hint: Compare to Theorem 6.4.)
- (12) Show that if W_1, W_2 are subspaces of V such that $W_1 \cap W_2 = \{\mathbf{0}\}$, then for any linearly independent vectors $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k$ in W_1 and $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_\ell$ in W_2 , the set $\{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k, \mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_\ell\}$ is linearly independent in V .
- (13) Label each of the following statements as True or False. Provide justification for your response. Throughout, let V be a vector space.
- True/False** If \mathbf{v} is in V , then the set $\{\mathbf{v}\}$ is linearly independent.
 - True/False** If a set of vectors span a subspace, then the set forms a basis of this subspace.
 - True/False** If a linearly independent set of vectors spans a subspace, then the set forms a basis of this subspace.
 - True/False** If the set S spans V and removing any vector from S makes it not a spanning set anymore, then S is a basis.
 - True/False** If S is a linearly independent set in V and for every \mathbf{u} in V , adding \mathbf{u} to S makes it not linearly independent anymore, then S is a basis.
 - True/False** If a subset S of V spans V , then S must be linearly independent.
 - True/False** If a subset S of V is linearly independent, then S must span V .
 - True/False** If S is a linearly dependent set in V , then every vector in S is a linear combination of the other vectors in S .
 - True/False** A vector space cannot have more than one basis.
 - True/False** If \mathbf{u} is a non-zero vector in V , then there is a basis of V containing \mathbf{u} .
 - True/False** If \mathbf{u}, \mathbf{v} are two linearly independent vectors in V , then there is a basis of V containing \mathbf{u}, \mathbf{v} .
 - True/False** If \mathbf{u} is in a basis of V , then $2\mathbf{u}$ cannot be in a basis of V .



Project: Image Compression with Wavelets

We return to the problem of image compression introduced at the beginning of this section. The first step in the wavelet compression process is to digitize an image. There are two important ideas about digitalization to understand here: intensity levels and resolution. In grayscale image processing, it is common to think of 256 different intensity levels, or scales, of gray ranging from 0 (black) to 255 (white). A digital image can be created by taking a small grid of squares (called pixels) and coloring each pixel with some shade of gray. The resolution of this grid is a measure of how many pixels are used per square inch. An example of a 16 by 16 pixel picture of a flower was shown in Figure 23.1.

An image can be thought of in several ways: as a two-dimensional array; as one long vector by stringing the columns together one after another; or as a collection of column vectors. For simplicity, we will use the last approach in this project. We call each column vector in a picture a *signal*. Wavelets are used to process signals. After processing we can apply some technique to compress the processed signals.

To process a signal we select a family of wavelets. There are many different families of wavelets – which family to use depends on the problem to be addressed. The simplest family of wavelets is the Haar family. More complicated families of wavelets are usually used in applications, but the basic ideas in wavelets can be seen through working with the Haar wavelets, and their relative simplicity will make the details easier to follow. Each family of wavelets has a father wavelet (usually denoted φ) and a mother wavelet (ψ).

Wavelets are generated from the mother wavelet by scalings and translations. To further simplify our work we will restrict ourselves to wavelets on $[0,1]$, although this is not necessary. The advantage the wavelets have over other methods of data analysis (Fourier analysis for example) is that with the scalings and translations we are able to analyze both frequency on large intervals and isolate signal discontinuities on very small intervals. The way this is done is by using a large collection (infinite, in fact) of basis functions with which to transform the data. We'll begin by looking at how these basis functions arise.

If we sample data at various points, we can consider our data to represent a piecewise constant function obtained by partitioning $[0,1]$ into n equal sized subintervals, where n represents the number of sample points. For the purposes of this project we will always choose n to be a power of 2. So we can consider all of our data to represent functions. For us, then, it is natural to look at these functions in the vector space of all functions from \mathbb{R} to \mathbb{R} . Since our data is piecewise constant, we can really restrict ourselves to a subspace of this larger vector space – subspaces of piecewise constant functions. The most basic piecewise constant function on the interval $[0, 1]$ is the one whose value is 1 on the entire interval. We define φ to be this constant function (called the characteristic function of the unit interval). That is

$$\varphi(x) = \begin{cases} 1 & \text{if } 0 \leq x < 1 \\ 0, & \text{otherwise.} \end{cases}$$

This function φ is the Father Haar wavelet.

This function φ may seem to be a very simple function but it has properties that will be important to us. One property is that φ satisfies a scaling equation. For example, Figure 23.2 shows



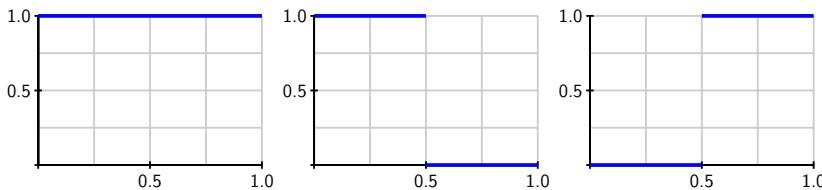


Figure 23.2: Graphs of $\varphi(x)$, $\varphi(2x)$, and $\varphi(2x - 1)$ from left to right.

that

$$\varphi(x) = \varphi(2x) + \varphi(2x - 1)$$

while Figure 23.3 shows that

$$\varphi(x) = \varphi(2^2x) + \varphi(2^2x - 1) + \varphi(2^2x - 2) + \varphi(2^2x - 3).$$

So φ is a sum of scalings and translations of itself. In general, for each positive integer n and

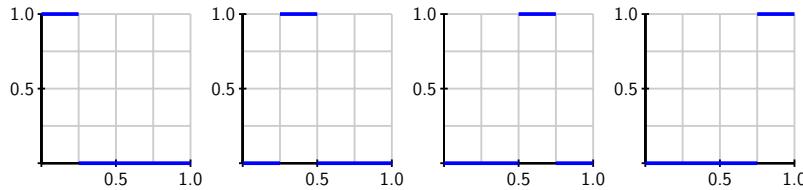


Figure 23.3: Graphs of $\varphi(2^2x)$, $\varphi(2^2x - 1)$, $\varphi(2^2x - 2)$, and $\varphi(2^2x - 3)$, from left to right.

integers k between 0 and $2^n - 1$ we define

$$\varphi_{n,k}(x) = \varphi(2^n x - k).$$

Then $\varphi(x) = \sum_{k=0}^{2^n-1} \varphi_{n,k}(x)$ for each n .

These functions $\varphi_{n,k}$ are useful in that they form a basis for the vector space V_n of all piecewise constant functions on $[0, 1]$ that have possible breaks at the points $\frac{1}{2^n}, \frac{2}{2^n}, \frac{3}{2^n}, \dots, \frac{2^n-1}{2^n}$. This is exactly the kind of space in which digital signals live, especially if we sample signals at 2^n evenly spaced points on $[0, 1]$. Let $\mathcal{B}_n = \{\varphi_{n,k} : 0 \leq k \leq 2^n - 1\}$. You may assume without proof that \mathcal{B}_n is a basis of V_n .

Project Activity 23.1.

- Draw the linear combination $2\varphi_{2,0} - 3\varphi_{2,1} + 17\varphi_{2,2} + 30\varphi_{2,3}$. What does this linear combination look like? Explain the statement made previously “Notice that these 2^n functions $\varphi_{n,k}$ form a basis for the vector space of all piecewise constant functions on $[0, 1]$ that have possible breaks at the points $\frac{1}{2^n}, \frac{2}{2^n}, \frac{3}{2^n}, \dots, \frac{2^n-1}{2^n}$ ”.
- Remember that we can consider our data to represent a piecewise constant function obtained by partitioning $[0, 1]$ into n subintervals, where n represents the number of sample points. Suppose we collect the following data: 10, 13, 21, 55, 3, 12, 4, 18. Explain how we can use this data to define a piecewise constant function f on $[0, 1]$. Express f as a linear combination of suitable functions $\varphi_{n,k}$. Plot this linear combination of $\varphi_{n,k}$ to verify.

Working with functions can be more cumbersome than working with vectors in \mathbb{R}^n , but the digital nature of our data makes it possible to view our piecewise constant functions as vectors in \mathbb{R}^n for suitable n . More specifically, if f is an element in V_n , then f is a piecewise constant function on $[0, 1]$ with possible breaks at the points $\frac{1}{2^n}, \frac{2}{2^n}, \frac{3}{2^n}, \dots, \frac{2^n-1}{2^n}$. If f has the value of y_i on the interval between $\frac{i-1}{2^n}$ and $\frac{i}{2^n}$, then we can identify f with the vector $[y_1 \ y_1 \ \dots \ y_{2^n}]^\top$.

Project Activity 23.2.

- (a) Determine the vector in \mathbb{R}^8 that is identified with φ .
- (b) Determine the value of m and the vectors in \mathbb{R}^m that are identified with $\varphi_{2,0}, \varphi_{2,1}, \varphi_{2,2}$, and $\varphi_{2,3}$.

We can use the functions $\varphi_{n,k}$ to represent digital signals, but to manipulate the data in useful ways we need a different perspective. A different basis for V_n (a *wavelet basis*) will allow us to identify the pieces of the data that are most important. We illustrate in the next activity with the spaces V_1 and V_2 .

Project Activity 23.3. The space V_1 consists of all functions that are piecewise constant on $[0, 1]$ with a possible break at $x = \frac{1}{2}$. The functions $\varphi = \varphi_{n,k}$ are used to records the values of a signal, and by summing these values we can calculate their average. Wavelets act by averaging and differencing, and so φ does the averaging. We need functions that will perform the differencing.

- (a) Define $\{\psi_{0,0}\}$ as

$$\psi_{0,0}(x) = \begin{cases} 1 & \text{if } 0 \leq x < \frac{1}{2} \\ -1 & \text{if } \frac{1}{2} \leq x < 1 \\ 0 & \text{otherwise} \end{cases}$$

A picture of $\psi_{0,0}$ is shown in Figure 23.4. Since $\psi_{0,0}$ assumes values of 1 and -1 , we can use $\psi_{0,0}$ to perform differencing. The function $\psi = \psi_{0,0}$ is the Mother Haar wavelet.³ Show that $\{\varphi, \psi\}$ is a basis for V_1 .

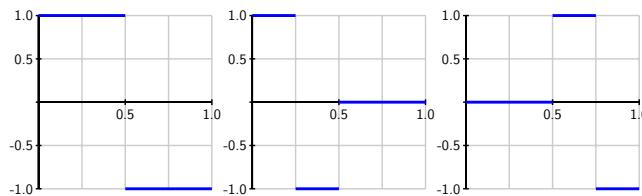


Figure 23.4: The graphs of $\psi_{0,0}, \psi_{1,0}$ and $\psi_{1,1}$ from left to right.

- (b) We continue in a manner similar to the one in which we constructed bases for V_n . For $k = 0$ and $k = 1$, let $\psi_{1,k} = \psi(2^1x - k)$. Graphs of $\psi_{1,0}$ and $\psi_{1,1}$ are shown in Figure 23.4. The functions $\psi_{1,k}$ assume the values of 1 and -1 on smaller intervals, and so can be used to perform differencing on smaller scale than $\psi_{0,0}$. Show that $\{\varphi_{0,0}, \psi_{0,0}, \psi_{1,0}, \psi_{1,1}\}$ is a basis for V_2 .

³The first mention of wavelets appeared in an appendix to the thesis of A. Haar in 1909.

As Project Activity 23.3 suggests, we can make a basis for V_n from $\varphi_{0,0}$ and functions of the form $\psi_{n,k}$ defined by $\psi_{n,k}(x) = \psi(2^n x - k)$ for k from 0 to $2^n - 1$. More specifically, if we let $\mathcal{S}_n = \{\psi_{n,k} : 0 \leq k \leq 2^n - 1\}$, then the set

$$\mathcal{W}_n = \{\varphi_{0,0}\} \cup \bigcup_{j=0}^{n-1} \mathcal{S}_j$$

is a basis for V_n^\perp (we state this without proof). The functions $\psi_{n,k}$ are the *wavelets*.

Project Activity 23.4. We can now write any function in V_n using the basis \mathcal{W}_n . As an example, the string 50, 16, 14, 28 represents a piecewise constant function which can be written as $50\varphi_{2,0} + 16\varphi_{2,1} + 14\varphi_{2,2} + 28\varphi_{2,3}$, an element in V_2 .

- (a) Specifically identify the functions in \mathcal{W}_0 , \mathcal{W}_1 , and \mathcal{W}_2 , and \mathcal{W}_3 .
- (b) As mentioned earlier, we can identify a signal, and each wavelet function, with a vector in \mathbb{R}^m for an appropriate value of m . We can then use this identification to decompose any signal as a linear combination of wavelets. We illustrate this idea with the signal $[50 \ 16 \ 14 \ 28]^\top$ in \mathbb{R}^4 . Recall that we can represent this signal as the function $f = 50\varphi_{2,0} + 16\varphi_{2,1} + 14\varphi_{2,2} + 28\varphi_{2,3}$.
 - i. Find the vectors \mathbf{w}_1 , \mathbf{w}_2 , \mathbf{w}_3 , and \mathbf{w}_4 in \mathbb{R}^m that are identified with $\varphi_{0,0}$, $\psi_{0,0}$, $\psi_{1,0}$, and $\psi_{1,1}$, respectively.
 - ii. Any linear combination $c_1\varphi_{0,0} + c_2\psi_{0,0} + c_3\psi_{1,0} + c_4\psi_{1,1}$ is then identified with the linear combination $c_1\mathbf{w}_1 + c_2\mathbf{w}_2 + c_3\mathbf{w}_3 + c_4\mathbf{w}_4$. Use this idea to find the weights to write the function f as a linear combination of $\varphi_{0,0}$, $\psi_{0,0}$, $\psi_{1,0}$, and $\psi_{1,1}$.

Although it is not necessarily easy to observe, the weights in the decomposition $f = 27\varphi_{0,0} + 6\psi_{0,0} + 17\psi_{1,0} - 7\psi_{1,1}$ are just averages and differences of the original weights in $f = 50\varphi_{2,0} + 16\varphi_{2,1} + 14\varphi_{2,2} + 28\varphi_{2,3}$. To see how, notice that if we take the overall average of the original weights we obtain the value of 27. If we average the original weights in pairs (50 and 16, and 14 and 28) we obtain the values 33 and 21, and if we take average differences of the original weights in pairs (50 and 16, and 14 and 28) we obtain the values 17 and -7 . We can treat the signal $[33 \ 21]^\top$ formed from the average of the pairs of the original weights as a smaller copy of the original signal. The average difference of the entries of this new signal is 6. So the weights in our final decomposition are obtained by differences between successive averages and certain coefficients. The coefficients in our final decomposition $27\varphi_{0,0} + 6\psi_{0,0} + 17\psi_{1,0} - 7\psi_{1,1}$ are called *wavelet coefficients*. This is the idea that makes wavelets so useful for image compression. In many images, pixels that are near to each other often have similar coloring or shading. These pixels are coded with numbers that are close in value. In the differencing process, these numbers are replaced with numbers that are close to 0. If there is little difference in the shading of the adjacent pixels, the image will be changed only a little if the shadings are made the same. This results in replacing these small wavelet coefficients with zeros. If the processed vectors contain long strings of zeros, the vectors can be significantly compressed.

Once we have recognized the pattern in expressing our original function as an overall average and wavelet coefficients we can perform these operations more quickly with matrices.



Project Activity 23.5. The process of averaging and differencing discussed in and following Project Activity 23.4 can be viewed as a matrix-vector problem. As we saw in Project Activity 23.4, we can translate the problem of finding wavelet coefficients to the matrix world.

- Consider again the problem of finding the wavelet coefficients contained in the vector $[27 \ 6 \ 17 \ -7]^T$ for the signal $[50 \ 16 \ 14 \ 28]^T$. Find the matrix A_4 that has the property that $A_4[50 \ 16 \ 14 \ 28]^T = [27 \ 6 \ 17 \ -7]^T$. (You have already done part of this problem in Project Activity 23.4.) Explain how A_4 performs the averaging and differencing discussed earlier.
- Repeat the process in part (a) to find the matrix A_8 that converts a signal to its wavelet coefficients.
- The matrix A_i is called a *forward wavelet transformation matrix* and A_i^{-1} is the *inverse wavelet transform matrix*. Use A_8 to show that the wavelet coefficients for the data string $[80 \ 48 \ 4 \ 36 \ 28 \ 64 \ 6 \ 50]^T$ are contained in the vector $[39.5 \ 2.5 \ 22 \ 9 \ 16 \ -16 \ -18 \ -22]^T$.

Now we have all of the necessary background to discuss image compression. Suppose we want to store an image. We partition the image vertically and horizontally and record the color or shade at each grid entry. The grid entries will be our pixels. This gives a matrix, M , of colors, indexed by pixels or horizontal and vertical position. To simplify our examples we will work in gray-scale, where our grid entries are integers between 0 (black) and 255 (white). We can treat each column of our grid as a piecewise constant function. As an example, the image matrix M that produced the picture at left in Figure 23.1 is given in (23.1).

We can then apply a 16 by 16 forward wavelet transformation matrix A_{16} to M to convert the columns to averages and wavelet coefficients that will appear in the matrix $A_{16}M$. These wavelet coefficients allow us to compress the image – that is, create a smaller set of data that contains the essence of the original image.

Recall that the forward wavelet transformation matrix computes weighted differences of consecutive entries in the columns of the image matrix M . If two entries in M are close in values, the weighted difference in $A_{16}M$ will be close to 0. For our example, the matrix $A_{16}M$ is approximately

$$\begin{bmatrix} 208.0 & 202.0 & 178.0 & 165.0 & 155.0 & 172.0 & 118.0 & 172.0 & 155.0 & 153.0 & 176.0 & 202.0 & 208.0 & 210.0 & 209.0 & 208.0 \\ 33.4 & 24.1 & -0.625 & 0.938 & -2.50 & -5.94 & 42.8 & -5.94 & -2.50 & 12.8 & 0.938 & 24.7 & 30.6 & 33.4 & 32.5 & 31.6 \\ -1.88 & -13.8 & 19.4 & 2.50 & 0.0 & -2.50 & 8.12 & -2.50 & 0.0 & 2.50 & 19.4 & -13.8 & 1.88 & -3.75 & -1.88 & 0.0 \\ 17.5 & 61.9 & 61.9 & 6.88 & 0.0 & 61.9 & 0.0 & 61.9 & 0.0 & 30.6 & 65.0 & 66.9 & 66.9 & 19.4 & 66.9 & 66.9 \\ 0.0 & 27.5 & 43.8 & 16.2 & 0.0 & -11.2 & 16.2 & -11.2 & 0.0 & 16.2 & 43.8 & 27.5 & 0.0 & 0.0 & 0.0 & 0.0 \\ 3.75 & 0.0 & 27.5 & -11.2 & 0.0 & -16.2 & 22.5 & -16.2 & 0.0 & -11.2 & 27.5 & 0.0 & -3.75 & -7.50 & -3.75 & 0.0 \\ 47.5 & 0.0 & 0.0 & 13.8 & 82.5 & 0.0 & 0.0 & 82.5 & 13.8 & 0.0 & 3.75 & 3.75 & 51.2 & 3.75 & 3.75 & 0.0 \\ 82.5 & 41.2 & 41.2 & 82.5 & 82.5 & 41.2 & 0.0 & 41.2 & 82.5 & 35.0 & 35.0 & 35.0 & 35.0 & 82.5 & 35.0 & 35.0 \\ 0.0 & 0.0 & 0.0 & 55.0 & -22.5 & -22.5 & 55.0 & -22.5 & -22.5 & 55.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 \\ 0.0 & 55.0 & -22.5 & -22.5 & 22.5 & 0.0 & -22.5 & 0.0 & 22.5 & -22.5 & -22.5 & 55.0 & 0.0 & 0.0 & 0.0 & 0.0 \\ -7.50 & 0.0 & -55.0 & 22.5 & 22.5 & -22.5 & 0.0 & -22.5 & 22.5 & 22.5 & -55.0 & 0.0 & 7.50 & 0.0 & 0.0 & 0.0 \\ 0.0 & 0.0 & 0.0 & 0.0 & 22.5 & -55.0 & 0.0 & -55.0 & 22.5 & 0.0 & 0.0 & 0.0 & -15.0 & 0.0 & -7.50 & 0.0 \\ 0.0 & 0.0 & 0.0 & -55.0 & 0.0 & 0.0 & 0.0 & 0.0 & -55.0 & 0.0 & 7.50 & 7.50 & 7.50 & 7.50 & 7.50 & 7.50 \\ 95.0 & 0.0 & 0.0 & -82.5 & 0.0 & 0.0 & 0.0 & 0.0 & -82.5 & 0.0 & 0.0 & 0.0 & 95.0 & 0.0 & 0.0 & 0.0 \\ 0.0 & -82.5 & 82.5 & 0.0 & 0.0 & -82.5 & 0.0 & -82.5 & 0.0 & 95.0 & -95.0 & 95.0 & -95.0 & 0.0 & 95.0 & -95.0 \\ 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 \end{bmatrix}.$$

Note that there are many wavelet coefficients that are quite small compared to others – the ones where the weighted averages are close to 0. In a sense, the weighted differences tell us how much “detail” about the whole that each piece of information contains. If a piece of information contains only a small amount of information about the whole, then we shouldn’t sacrifice much of the picture if we ignore the small “detail” coefficients. One way to ignore the small “detail” coefficients is to use *thresholding*.

With thresholding (this is *hard thresholding* or *keep or kill*), we decide on how much of the detail we want to remove (this is called the *tolerance*). So we set a tolerance and then replace each entry in our matrix $A_{16}M$ whose absolute value is below the tolerance with 0 to obtain a new matrix M_1 . In our example, if you use a threshold value of 10 we obtain the new matrix M_1 :

$$\begin{bmatrix} 208.0 & 202.0 & 178.0 & 165.0 & 155.0 & 172.0 & 118.0 & 172.0 & 155.0 & 153.0 & 176.0 & 202.0 & 208.0 & 210.0 & 209.0 & 208.0 \\ 33.4 & 24.1 & 0.0 & 0.0 & 0.0 & 0.0 & 42.8 & 0.0 & 0.0 & 12.8 & 0.0 & 24.7 & 30.6 & 33.4 & 32.5 & 31.6 \\ 0.0 & -13.8 & 19.4 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 19.4 & -13.8 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 \\ 17.5 & 61.9 & 61.9 & 0.0 & 0.0 & 61.9 & 0.0 & 61.9 & 0.0 & 30.6 & 65.0 & 66.9 & 66.9 & 19.4 & 66.9 & 66.9 \\ 0.0 & 27.5 & 43.8 & 16.2 & 0.0 & -11.2 & 16.2 & -11.2 & 0.0 & 16.2 & 43.8 & 27.5 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 \\ 0.0 & 0.0 & 27.5 & -11.2 & 0.0 & -16.2 & 22.5 & -16.2 & 0.0 & -11.2 & 27.5 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 \\ 47.5 & 0.0 & 0.0 & 13.8 & 82.5 & 0.0 & 0.0 & 82.5 & 0.0 & 13.8 & 0.0 & 0.0 & 0.0 & 51.2 & 0.0 & 0.0 & 0.0 \\ 82.5 & 41.2 & 41.2 & 82.5 & 82.5 & 41.2 & 0.0 & 41.2 & 82.5 & 35.0 & 35.0 & 35.0 & 35.0 & 82.5 & 35.0 & 35.0 & 35.0 \\ 0.0 & 0.0 & 0.0 & 55.0 & -22.5 & -22.5 & 55.0 & -22.5 & -22.5 & 55.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 \\ 0.0 & 55.0 & -22.5 & -22.5 & 22.5 & 0.0 & -22.5 & 0.0 & 22.5 & -22.5 & -22.5 & 55.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 \\ 0.0 & 0.0 & -55.0 & 22.5 & 22.5 & -22.5 & 0.0 & -22.5 & 22.5 & 22.5 & -55.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 \\ 0.0 & 0.0 & 0.0 & 0.0 & 22.5 & -55.0 & 0.0 & -55.0 & 22.5 & 0.0 & 0.0 & 0.0 & -15.0 & 0.0 & 0.0 & 0.0 & 0.0 \\ 0.0 & 0.0 & 0.0 & -55.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & -55.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 \\ 95.0 & 0.0 & 0.0 & -82.5 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & -82.5 & 0.0 & 0.0 & 0.0 & 95.0 & 0.0 & 0.0 & 0.0 \\ 0.0 & -82.5 & 82.5 & 0.0 & 0.0 & -82.5 & 0.0 & -82.5 & 0.0 & 95.0 & -95.0 & 95.0 & -95.0 & 0.0 & 95.0 & 0.0 & -95.0 \\ 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 \end{bmatrix}.$$

We now have introduced many zeros in our matrix. This is where we compress the image. To store the original image, we need to store every pixel. Once we introduce strings of zeros we can identify a new code (say 256) that indicates we have a string of zeros. We can then follow that code with the number of zeros in the string. So if we had a string of 15 zeros in a signal, we could store that information in 2 bytes rather than 15 and obtain significant savings in storage. This process removes some detail from our picture, but only the small detail. To convert back to an image, we just undo the forward processing by multiplying our thresholded matrix M_1 by A_{16}^{-1} . The ultimate goal is to obtain significant compression but still have $A_{16}^{-1}M_1$ retain all of the essence of the original image.

In our example using M_1 , the reconstructed image matrix is $A_{16}^{-1}M_1$ (rounded to the nearest whole number) is

$$\begin{bmatrix} 242 & 240 & 241 & 237 & 132 & 138 & 232 & 138 & 132 & 238 & 239 & 240 & 238 & 244 & 242 & 240 \\ 242 & 240 & 241 & 127 & 178 & 183 & 122 & 183 & 178 & 128 & 239 & 240 & 238 & 244 & 242 & 240 \\ 242 & 240 & 131 & 127 & 178 & 183 & 122 & 183 & 178 & 128 & 129 & 240 & 238 & 244 & 242 & 240 \\ 242 & 130 & 176 & 172 & 132 & 183 & 167 & 183 & 132 & 172 & 174 & 130 & 238 & 244 & 242 & 240 \\ 242 & 240 & 131 & 177 & 178 & 133 & 183 & 133 & 178 & 178 & 129 & 240 & 238 & 244 & 242 & 240 \\ 242 & 240 & 241 & 132 & 132 & 178 & 183 & 178 & 132 & 132 & 239 & 240 & 238 & 244 & 242 & 240 \\ 242 & 240 & 131 & 177 & 178 & 133 & 138 & 133 & 178 & 178 & 129 & 240 & 223 & 244 & 242 & 240 \\ 242 & 240 & 131 & 177 & 132 & 243 & 138 & 243 & 132 & 178 & 129 & 240 & 253 & 244 & 242 & 240 \\ 240 & 240 & 239 & 124 & 238 & 234 & 75 & 234 & 238 & 130 & 241 & 244 & 244 & 248 & 244 & 244 \\ 240 & 240 & 239 & 234 & 238 & 234 & 75 & 234 & 238 & 240 & 241 & 244 & 244 & 248 & 244 & 244 \\ 240 & 240 & 239 & 69 & 73 & 234 & 75 & 234 & 73 & 75 & 241 & 244 & 244 & 240 & 244 & 244 \\ 50 & 240 & 239 & 234 & 73 & 234 & 75 & 234 & 73 & 240 & 241 & 244 & 244 & 50 & 244 & 244 \\ 240 & 75 & 239 & 248 & 238 & 69 & 75 & 69 & 238 & 240 & 51 & 240 & 50 & 240 & 240 & 50 \\ 240 & 240 & 74 & 248 & 238 & 234 & 75 & 234 & 238 & 50 & 241 & 50 & 240 & 240 & 50 & 240 \\ 75 & 75 & 74 & 83 & 73 & 69 & 75 & 69 & 73 & 75 & 76 & 75 & 75 & 75 & 75 & 75 \\ 75 & 75 & 74 & 83 & 73 & 69 & 75 & 69 & 73 & 75 & 76 & 75 & 75 & 75 & 75 & 75 \end{bmatrix}.$$

We convert this into a gray-scale image and obtain the image at right in Figure 23.1. Compare this image to the original at right in Figure 23.1. It is difficult to tell the difference.

There is a Sage file you can use at



http://faculty.gvsu.edu/schlicks/Wavelets_Sage.html

that allows you to create your own 16 by 16 image and process, process your image with the Haar wavelets in \mathbb{R}^{16} , apply thresholding, and reconstruct the compressed image. matrix. You can create your own image, experiment with several different threshold levels, and choose the one that you feel gives the best combination of strings of 0s while reproducing a reasonable copy of the original image.



Section 24

The Dimension of a Vector Space

Focus Questions

By the end of this section, you should be able to give precise and thorough answers to the questions listed below. You may want to keep these questions in mind to focus your thoughts as you complete the section.

- What is a finite dimensional vector space?
- What is the dimension of a finite dimensional vector space? What important result about bases of finite dimensional vector spaces makes dimension well-defined?
- What must be true about any linearly independent subset of n vectors in a vector space with dimension n ? Why?
- What must be true about any subset of n vectors in a vector space with dimension n that spans the vector space? Why?

Application: Principal Component Analysis

The discipline of statistics is based on the idea of analyzing data. In large data sets it is usually the case that one wants to understand the relationships between the different data in the set. This can be difficult to do when the data set is large and it is impossible to visually examine the data for patterns. Principal Component Analysis (PCA) is a tool for identifying and representing underlying patterns in large data sets, and PCA has been called one of the most important and valuable linear algebra tools for statistical analysis. PCA is used to transform a collection of variables into a (usually smaller) number of uncorrelated variables called principal components. The principal components form the most meaningful basis from which to view data by removing extraneous information and revealing underlying relationships in the data. This presents a framework for how to reduce a complex data set to a lower dimension while retaining the important attributes of the data set. The output helps the experimenter determine which dynamics in the data are important and which can be ignored.

Introduction

In Section 15 we learned that any two bases for a subspace of \mathbb{R}^n contain the same number of vectors. This allowed us to define the dimension of a subspace of \mathbb{R}^n . In this section we extend the arguments we made in Section 15 to arbitrary vector spaces and define the dimension of a vector space.

Preview Activity 24.1. The main tool we used to prove that any two bases for a subspace of \mathbb{R}^n must contain the same number of elements was Theorem 15.1. In this preview activity we will show that the same argument can be used for vector spaces. More specifically, we will prove a special case of the following theorem generalizing Theorem 15.1.

Theorem 24.1. *If V is a vector space with a basis $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k\}$ of k vectors, then any subset of V containing more than k vectors is linearly dependent.*

Suppose V is a vector space with basis $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2\}$. Consider the set $U = \{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ of vectors in V . We will show that U is linearly dependent using a similar approach to the Preview Activity 15.1.

- (1) What vector equation involving $\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3$ do we need to solve to determine linear independence/dependence of these vectors? Use x_1, x_2, x_3 for coefficients.
- (2) Since \mathcal{B} is a basis of V , it spans V . Using this information, rewrite the vectors \mathbf{u}_i in terms of \mathbf{v}_j and substitute into the above equation to obtain another equation in terms of \mathbf{v}_j .
- (3) Since \mathcal{B} is a basis of V , the vectors $\mathbf{v}_1, \mathbf{v}_2$ are linearly independent. Using the equation in the previous part, determine what this means about the coefficients x_1, x_2, x_3 .
- (4) Express the conditions on x_1, x_2, x_3 in the form of a matrix-vector equation. Explain why there are infinitely many solutions for x_i 's and why this means the vectors $\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3$ are linearly dependent.

Finite Dimensional Vector Spaces

Theorem 24.1 shows that if sets \mathcal{B}_1 and \mathcal{B}_2 are finite bases for a vector space V , which are linearly independent by definition, then each cannot contain more elements than the other, so the number of elements in each basis must be equal.

Theorem 24.2. *If a non-trivial vector space V has a basis of n vectors, then every basis of V contains exactly n vectors.*

Theorem 24.2 states that if a vector space V has a basis with a finite number of vectors, then the number of vectors in a basis for that vector space is a well-defined number. In other words, the number of vectors in a basis is an *invariant* of the vector space. This important number is given a name.

Definition 24.3. A **finite-dimensional** vector space is a vector space that can be spanned by a finite number of vectors. The **dimension** of a non-trivial finite-dimensional vector space is the number of vectors in a basis for V . The dimension of the trivial vector space is defined to be 0.



We denote the dimension of a finite dimensional vector space V by $\dim(V)$.

Not every vector space is finite dimensional. We have seen, for example, that the vector space \mathbb{P} of all polynomials, regardless of degree, is not a finite-dimensional vector space. In fact, the polynomials

$$1, t, t^2, \dots, t^n, \dots$$

are linearly independent, so \mathbb{P} has an infinite linearly independent set and therefore has no finite basis. A vector space that has an infinite basis is called an *infinite dimensional* vector space.

Activity 24.1. Since columns of the $n \times n$ identity matrix span \mathbb{R}^n and are linearly independent, the columns of I_n form a basis for \mathbb{R}^n (the standard basis). Consequently, we have that $\dim(\mathbb{R}^n) = n$. In this activity we determine the dimensions of other familiar vector spaces. Find the dimensions of each of the indicated vector spaces. Verify your answers.

- | | | |
|--------------------------------|--------------------------------|--------------------------------|
| (a) \mathbb{P}_1 | (b) \mathbb{P}_2 | (c) \mathbb{P}_n |
| (d) $\mathcal{M}_{2 \times 3}$ | (e) $\mathcal{M}_{3 \times 4}$ | (f) $\mathcal{M}_{k \times n}$ |

Finding the dimension of a finite-dimensional vector space amounts to finding a basis for the space.

Activity 24.2. Let $W = \{(a + b) + (a - b)t + (2a + 3b)t^2 \mid a, b \text{ are scalars}\}$.

- (a) Find a finite set of polynomials in W that span W .
- (b) Determine if the spanning set from part (a) is linearly independent or dependent. Clearly explain your process.
- (c) What is $\dim(W)$? Explain.

The Dimension of a Subspace

Every subspace of a finite-dimensional vector space is a vector space, and since a subspace is contained in a vector space it is natural to think that the dimension of a subspace should be less than or equal to the dimension of the larger vector space. We verify that fact in this section.

Activity 24.3. Let V be a finite dimensional vector space of dimension n and let W be a subspace of V . Explain why W cannot have dimension larger than $\dim(V)$, and if $W \neq V$ then $\dim(W) < \dim(V)$. (Hint: Use Theorem 24.1.)

Conditions for a Basis of a Vector Space

There are two items we need to confirm before we can state that a subset B of a subspace W of a vector space is a basis for W : the set B must be linearly independent and span W . We can reduce the amount of work it takes to show that a set is a basis if we know the dimension of the vector space in advance.

Activity 24.4. Let W be a subspace of a vector space V with $\dim(W) = k$. We know that every basis of W contains exactly k vectors.



- (a) Suppose that S is a subset of W that contains k vectors and is linearly independent. In this part of the activity we will show that S must span W .
- Suppose that S does not span W . Explain why this implies that W contains a set of $k + 1$ linearly independent vectors.
 - Explain why the result of i tells us that S is a basis for W .
- (b) Now suppose that S is a subset of W with k vectors that spans W . In this part of the activity we will show that S must be linearly independent.
- Suppose that S is not linearly independent. Explain why we can then find a proper subset of S that is linearly independent but has the same span as S .
 - Explain why the result of i tells us that S is a basis for W .

The result of Activity 24.4 is summarized in the following theorem (compare to Theorem 15.4).

Theorem 24.4. *Let W be a subspace of dimension k of a vector space V and let S be a subset of W containing exactly k vectors.*

- (1) *If S is linearly independent, then S is a basis for W .*
- (2) *If S spans W , then S is a basis for W .*

Examples

What follows are worked examples that use the concepts from this section.

Example 24.5. Find a basis and dimension for each of the indicated subspaces of the given vector spaces.

- $\{a + b(t + t^2) : a, b \in \mathbb{R}\}$ in \mathbb{P}_2
- $\text{Span} \left\{ 1, \frac{1}{1+x^2}, \frac{2+x^2}{1+x^2}, \arctan(x) \right\}$ in \mathcal{F}
- $\{p(t) \in \mathbb{P}_n : p(-t) = p(t)\}$ in \mathbb{P}_4 (The polynomials with the property that $p(-t) = p(t)$ are called *even* polynomials.)

Example Solution.

- Let $W = \{a + b(t + t^2) : a, b \in \mathbb{R}\}$. Every element in W has the form

$$a + b(t + t^2) = a(1) + b(t + t^2).$$

So $W = \text{Span}\{1, t + t^2\}$. Since neither 1 nor $t + t^2$ is a scalar multiple of the other, the set $\{1, t + t^2\}$ is linearly independent. Thus, $\{1, t + t^2\}$ is a basis for W and $\dim(W) = 2$.



- (b) Let $W = \text{Span} \left\{ 1, \frac{1}{1+x^2}, \frac{2+x^2}{1+x^2}, \arctan(x) \right\}$. To find a basis for W , we find a linearly independent subset of $\left\{ 1, \frac{1}{1+x^2}, \frac{2+x^2}{1+x^2}, \arctan(x) \right\}$. Consider the equation

$$c_1(1) + c_2 \left(\frac{1}{1+x^2} \right) + c_3 \left(\frac{2+x^2}{1+x^2} \right) + c_4 \arctan(x) = 0.$$

To find the weights c_i for which this equality of functions holds, we use the fact that the we must have equality for every x . So we pick four different values for x to obtain a linear system that we can solve for the weights. Evaluating both sides of the equation at $x = 0$, $x = 1$, $x = -1$, and $x = 2$ yields the equations

$$\begin{aligned} c_1 + c_2 + 2c_3 &= 0 \\ c_1 + \frac{1}{2}c_2 + \frac{3}{2}c_3 + &\quad \frac{\pi}{4}c_4 = 0 \\ c_1 + \frac{1}{2}c_2 + \frac{3}{2}c_3 - &\quad \frac{\pi}{4}c_4 = 0 \\ c_1 + \frac{1}{5}c_2 + \frac{6}{5}c_3 + \arctan(2)c_4 &= 0. \end{aligned}$$

The reduced row echelon form of the coefficient matrix

$$\left[\begin{array}{cccc} 1 & 1 & 2 & 0 \\ 1 & \frac{1}{2} & \frac{3}{2} & \frac{\pi}{4} \\ 1 & \frac{1}{2} & \frac{3}{2} & -\frac{\pi}{4} \\ 1 & \frac{1}{5} & \frac{6}{5} & \arctan(2) \end{array} \right]$$

is $\left[\begin{array}{cccc} 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{array} \right]$. The general solution to this linear system is $c_1 = c_2 = -c_3$ and $c_4 = 0$. Notice that

$$(-1)(1) + (-1) \left(\frac{1}{1+x^2} \right) + \frac{2+x^2}{1+x^2} = 0$$

or

$$\frac{2+x^2}{1+x^2} = 1 + \left(\frac{1}{1+x^2} \right),$$

so $\frac{2+x^2}{1+x^2}$ is a linear combination of the other vectors. The vectors corresponding to the pivot columns are linearly independent, so it follows that 1 , $\frac{1}{1+x^2}$, and $\arctan(x)$ are linearly independent. We conclude that $\left\{ 1, \frac{1}{1+x^2}, \arctan(x) \right\}$ is a basis for W and $\dim(W) = 3$.

- (c) Let $W = \{p(t) \in \mathbb{P}_n : p(-t) = p(t)\}$ in \mathbb{P}_4 . Let $p(t) \in W$ and suppose that $p(t) = a_0 + a_1 t + a_2 t^2 + a_3 t^3 + a_4 t^4$. Since $p(-t) = p(t)$, we must have $p(1) = p(-1)$ and $p(3) = p(-3)$. Since $p(1) = p(-1)$, it follows that

$$a_0 + a_1 + a_2 + a_3 + a_4 = a_0 - a_1 + a_2 - a_3 + a_4$$



or

$$a_1 + a_3 = 0.$$

Similarly, the fact that $p(3) = p(-3)$ yields the equation

$$a_0 + 3a_1 + 9a_2 + 27a_3 + 81a_4 = a_0 - 3a_1 + 9a_2 - 27a_3 + 81a_4$$

or

$$a_1 + 9a_3 = 0.$$

The reduced row echelon form of the coefficient matrix of system $a_1 + a_3 = 0$ and $a_1 + 9a_3 = 0$ is I_2 , so it follows that $a_1 = a_3 = 0$. Thus, $p(t) = a_0 + a_2t^2 + a_4t^4$ and so $W = \text{Span}\{1, t^2, t^4\}$. Equating like terms in the equation

$$c_1(1) + c_2(t^2) + c_3(t^4) = 0$$

yields $c_1 = c_2 = c_3 = 0$. We conclude that $\{1, t^2, t^4\}$ is linearly independent and is therefore a basis for W . Thus, $\dim(W) = 3$.

As an alternative solution, notice that $p(t) = t$ is not in W . So $W \neq V$ and we know that $\dim(W) < \dim(V)$. Since $1, t^2$, and t^4 are in W , we can show as above that $1, t^2$, and t^4 are linearly independent. We can conclude that $\dim(W) = 3$ since it cannot be 4.

Example 24.6. Let $U_{n \times n}$ be the set of all $n \times n$ upper triangular matrices. Recall that a matrix $A = [a_{ij}]$ is upper triangular if $a_{ij} = 0$ whenever $i > j$. That is, a matrix is upper triangular if all entries below the diagonal are 0.

- (a) Show that $U_{n \times n}$ is a subspace of $\mathcal{M}_{n \times n}$.
- (b) Find the dimensions of $U_{2 \times 2}$ and $U_{3 \times 3}$. Explain. Make a conjecture as to what $\dim(U_{n \times n})$ is in terms of n .

Example Solution.

- (a) Since the $n \times n$ zero matrix $0_{n \times n}$ has all entries equal to 0, it follows that $0_{n \times n}$ is in $U_{n \times n}$. Let $A = [a_{ij}]$ and $B = [b_{ij}]$ be in $U_{n \times n}$, and let $C = [c_{ij}] = A + B$. Then $c_{ij} = a_{ij} + b_{ij} = 0 + 0$ when $i > j$. So C is an upper triangular matrix and $U_{n \times n}$ is closed under addition. Let c be a scalar. The ij th entry of cA is $ca_{ij} = c(0) = 0$ whenever $i > j$. So cA is an upper triangular matrix and $U_{n \times n}$ is closed under multiplication by scalars. We conclude that $U_{n \times n}$ is a subspace of $\mathcal{M}_{n \times n}$.
- (b) Let $M_{11} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$, $M_{12} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$, and $M_{22} = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$. We will show that $S = \{M_{11}, M_{12}, M_{22}\}$ is a basis for $U_{2 \times 2}$. Consider the equation

$$x_1 M_{11} + x_2 M_{12} + x_3 M_{22} = 0.$$

Equating like entries shows that $x_1 = x_2 = x_3 = 0$, and so S is linearly independent. If $\begin{bmatrix} a & b \\ 0 & c \end{bmatrix}$ is in $U_{2 \times 2}$, then

$$\begin{bmatrix} a & b \\ 0 & c \end{bmatrix} = aM_{11} + bM_{12} + cM_{22}$$



and so S spans $U_{2 \times 2}$. Thus, S is a basis for $U_{2 \times 2}$ and so $\dim(U_{2 \times 2}) = 3$.

Similarly, for the 3×3 case let M_{ij} for $i \leq j$ be the 3×3 matrix with a 1 in the ij position and 0 in every other position. Let

$$S = \{M_{11}, M_{12}, M_{13}, M_{22}, M_{23}, M_{33}\}.$$

Equating corresponding entries shows that if

$$x_1M_{11} + x_2M_{12} + x_3M_{13} + x_4M_{22} + x_5M_{23} + x_6M_{33} = 0,$$

then $x_1 = x_2 = x_3 = x_4 = x_5 = x_6 = 0$. So S is a linearly independent set. If $A = [a_{ij}]$ is in $U_{3 \times 3}$, then $A = \sum_{i \geq j} a_{ij}M_{ij}$ and S spans $U_{3 \times 3}$. We conclude that S is a basis for $U_{3 \times 3}$ and $\dim(U_{3 \times 3}) = 6$.

In general, for an $n \times n$ matrix, the set of matrices M_{ij} , one for each entry on and above the diagonal, is a basis for $U_{n \times n}$. There are n such matrices for the entries on the diagonal. The number of entries above the diagonal is equal to half the total number of entries (n^2) minus half the number of entries on the diagonal (n). So there is a total of $\frac{n^2-n}{2}$ such matrices for the entries above the diagonal. Therefore,

$$\dim(U_{n \times n}) = n + \frac{n^2-n}{2} = \frac{n^2+n}{2}.$$

Summary

- A finite dimensional vector space is a vector space that can be spanned by a finite set of vectors.
- We showed that any two bases for a finite dimensional vector space *must* contain the same number of vectors. Therefore, we can define the *dimension* of a finite dimensional vector space V to be the number of vectors in any basis for V .
- If V is a vector space with dimension n and S is any linearly independent subset of V with n vectors, then S is a basis for V . Otherwise, we could add vectors to S to make a basis for V and then V would have a basis of more than n vectors.
- If V is a vector space with dimension n and S is any subset of V with n vectors that spans V , then S is a basis for V . Otherwise, we could remove vectors from S to obtain a basis for V and then V would have a basis of fewer than n vectors.
- For any finite dimensional space V and a subspace W of V , $\dim(W) \leq \dim(V)$.

Exercises

- (1) Let $W = \text{Span}\{1 + t^2, 2 + t + 2t^2 + t^3, 1 + t + t^3, t - t^2 + t^3\}$ in \mathbb{P}_3 . Find a basis for W . What is the dimension of W ?

- (2) Let $A = \begin{bmatrix} 1 & 2 \\ 1 & 0 \end{bmatrix}$ and $B = \begin{bmatrix} 1 & 0 \\ 1 & -1 \end{bmatrix}$.



- (a) Are A and B linearly independent or dependent? Verify your result.
- (b) Extend the set $S = \{A, B\}$ to a basis for $\mathcal{M}_{2 \times 2}$. That is, find a basis for $\mathcal{M}_{2 \times 2}$ that contains both A and B .
- (3) Let $A = \begin{bmatrix} 1 & 2 & 0 \\ 3 & 0 & 2 \end{bmatrix}$, $B = \begin{bmatrix} -1 & 1 & 1 \\ 2 & 4 & 0 \end{bmatrix}$, $C = \begin{bmatrix} 5 & 1 & -3 \\ 0 & -12 & 4 \end{bmatrix}$, $D = \begin{bmatrix} 5 & 4 & -2 \\ 5 & -8 & 6 \end{bmatrix}$, and $E = \begin{bmatrix} 2 & 0 & 0 \\ -2 & 0 & 1 \end{bmatrix}$ in $\mathcal{M}_{2 \times 3}$ and let $S = \{A, B, C, D, E\}$.
- (a) Is S a basis for $\mathcal{M}_{2 \times 3}$? Explain.
- (b) Determine if S is a linearly independent or dependent set. Verify your result.
- (c) Find a basis \mathcal{B} for $\text{Span } S$ that is a subset of S and write all of the vectors in S as linear combinations of the vectors in \mathcal{B} .
- (d) Extend your basis \mathcal{B} from part (c) to a basis $\mathcal{M}_{2 \times 3}$. Explain your method.
- (4) Determine the dimension of each of the following vector spaces.
- (a) $\text{Span}\{2, 1 + t, t^2\}$ in \mathbb{P}_2
- (b) The space of all polynomials in \mathbb{P}_3 whose constant terms is 0.
- (c) $\text{Nul } \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}$
- (d) $\text{Span}\{[1 \ 2 \ 0 \ 1 \ -1]^T, [0 \ 1 \ 1 \ 1 \ 0]^T, [0 \ 3 \ 2 \ 3 \ 0]^T, [-1 \ 0 \ 1 \ 1 \ 1]^T\}$ in \mathbb{R}^5
- (5) Let W be the set of matrices in $\mathcal{M}_{2 \times 2}$ whose diagonal entries sum to 0. Show that W is a subspace of $\mathcal{M}_{2 \times 2}$, find a basis for W , and then find $\dim(W)$.
- (6) Show that if W is a subspace of a finite dimensional vector space V , then any basis of W can be extended to a basis of V .
- (7) Let W be the set of all polynomials $a + bt + ct^2$ in \mathbb{P}_2 such that $a + b + c = 0$. Show that W is a subspace of \mathbb{P}_2 , find a basis for W , and then find $\dim(W)$.
- (8) Suppose W_1, W_2 are subspaces in a finite-dimensional space V .
- (a) Show that it is not true in general that $\dim(W_1) + \dim(W_2) \leq \dim(V)$.
- (b) Are there any conditions on W_1 and W_2 that will ensure that $\dim(W_1) + \dim(W_2) \leq \dim(V)$? (Hint: See problem 12 in the previous section.)
- (9) Suppose $W_1 \subseteq W_2$ are two subspaces of a finite-dimensional space. Show that if $\dim(W_1) = \dim(W_2)$, then $W_1 = W_2$.
- (10) Suppose W_1, W_2 are both three-dimensional subspaces inside \mathbb{R}^4 . In this exercise we will show that $W_1 \cap W_2$ contains a plane. Let $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ be a basis for W_1 and let $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ be a basis for W_2 .
- (a) If $\mathbf{v}_1, \mathbf{v}_2$, and \mathbf{v}_3 are all in W_1 , explain why $W_1 \cap W_2$ must contain a plane.
- (b) Now we consider the case where not all of $\mathbf{v}_1, \mathbf{v}_2$, and \mathbf{v}_3 are in W_1 . Since the arguments will be the same, let us assume that \mathbf{v}_1 is not in W_1 .

- i. Explain why the set $\mathcal{S} = \{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3, \mathbf{v}_1\}$ is a basis for \mathbb{R}^4 .
- ii. Explain why \mathbf{v}_2 and \mathbf{v}_3 can be written as linear combinations of the vectors in \mathcal{S} . Use these linear combinations to find two vectors that are in $W_1 \cap W_2$. Then show that these vectors span a plane in $W_1 \cap W_2$.
- (11) Label each of the following statements as True or False. Provide justification for your response.
- True/False** The dimension of a finite dimensional vector space is the minimum number of vectors needed to span that space.
 - True/False** The dimension of a finite dimensional vector space is the maximum number of linearly independent vectors that can exist in that space.
 - True/False** If n vectors span an n -dimensional vector space V , then these vectors form a basis of V .
 - True/False** Any set of n vectors form a basis in an n -dimensional vector space.
 - True/False** Every vector in a vector space V spans a one-dimensional subspace of V .
 - True/False** Any set of n linearly independent vectors in a vector space V of dimensional n is a basis for V .
 - True/False** If $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k\}$ is linearly independent in V , then $\dim(V) \geq k$.
 - True/False** If a set of k vectors span V , then any set of more than k vectors in V is linearly dependent.
 - True/False** If an infinite set of vectors span V , then V is infinite-dimensional.
 - True/False** If W_1, W_2 are both two-dimensional subspaces of \mathbb{R}^3 , then $W_1 \cap W_2 \neq \{0\}$.
 - True/False** If $\dim(V) = n$ and W is a subspace of V with dimension n , then $W = V$.

Project: Understanding Principal Component Analysis

Suppose we were to film an experiment involving a ball that is bouncing up and down. Naively, we set up several cameras to follow the process of the experiment from different perspectives and collect the data. All of this data tells us something about the bouncing ball, but there may be no perspective that tells us the most important piece of information – that axis along which the ball bounces. The question, then, is how we can extract from the data this most important piece of information. Principal Component Analysis (PCA) is a tool for just this type of analysis.

We will use an example to illustrate important concepts we will need. To realistically apply PCA we will have much more data than this, but for now we will restrict ourselves to only two variables so that we can visualize our results. Table 24.1 presents information from ten states on two attributes related to the SAT – Evidence-Based Reading and Writing (EBRW) score and Math score. The SAT is made up of three sections: Reading, Writing and Language (also just called



State	1	2	3	4	5	6	7	8	9	10
EBRW	595	540	522	508	565	512	643	574	534	539
Math	571	536	493	493	554	501	655	566	534	539

Table 24.1: SAT data.

Writing), and Math. The EBRW score is calculated by combining the Reading and Writing section scores – both the Math and EBRW are scored on a scale of 200-800.

Each attribute (Math, EBRW score) creates a vector whose entries are the student responses for that attribute. The data provides the average scores from participating students in each state. In this example we have two attribute vectors:

$$\mathbf{x}_1 = [595 \ 540 \ 522 \ 508 \ 565 \ 512 \ 643 \ 574 \ 534 \ 539]^T \text{ and}$$

$$\mathbf{x}_2 = [571 \ 536 \ 493 \ 493 \ 554 \ 501 \ 655 \ 566 \ 534 \ 539]^T.$$

These vectors form the rows of a 2×10 matrix

$$X_0 = \begin{bmatrix} \mathbf{x}_1^T \\ \mathbf{x}_2^T \end{bmatrix} = \begin{bmatrix} 595 & 540 & 522 & 508 & 565 & 512 & 643 & 574 & 534 & 539 \\ 571 & 536 & 493 & 493 & 554 & 501 & 655 & 566 & 534 & 539 \end{bmatrix}$$

that makes up our data set. A plot of the data is shown at left in Figure 24.1, where the EBRW score is along the horizontal axis and the math score is along the vertical axis. The question we want to

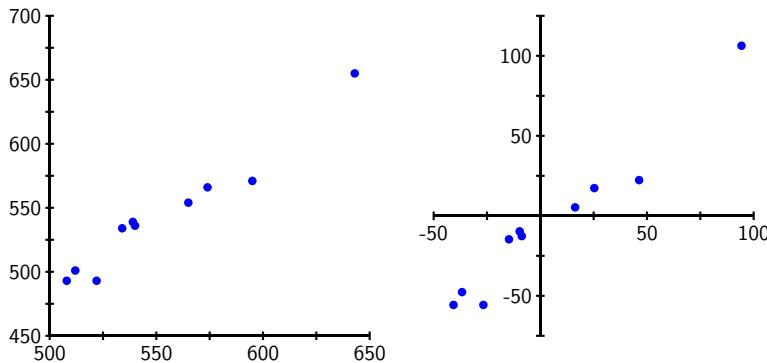


Figure 24.1: Two views of the data set (EBRW horizontal, math vertical).

answer is, how do we represent our data set so that the most important features in the data set are revealed?

Project Activity 24.1. Before analyzing a data set there is often some initial preparation that needs to be made. Issues that have to be dealt with include the problem that attributes might have different units and scales. For example, in a data set with attributes about people, height could be measured in inches while weight is measured in pounds. It is difficult to compare variables when they are on different scales. Another issue to consider is that some attributes are independent of the others (height, for example does not depend on hair color), while some are interrelated (body mass index depends on height and weight). To simplify our work, we will not address these type of problems. The only preparation we will do with our data is to center it.

- (a) An important piece of information about a one-dimensional data set $\mathbf{x} = [x_1 \ x_2 \ x_3 \ \cdots \ x_n]^T$ is the sample average or mean

$$\bar{\mathbf{x}} = \sum_{i=1}^n x_i.$$

Calculate the means \bar{x}_1 and \bar{x}_2 for our SAT data from the matrix X_0 .

- (b) We can use these sample means to center our data at the origin by translating the data so that each column of our data matrix has mean 0. We do this by subtracting the mean for that row vector from each component of the vector. Determine the matrix X that contains the centered data for our SAT data set from matrix X_0 .

A plot of the centered data for our SAT data is shown at right in Figure 24.1. Later we will see why centering the data is useful – it will more easily allow us to project onto subspaces. The goal of PCA is to find a matrix P so that $PX = Y$, and Y is suitably transformed to identify the important aspects of the data. We will discuss what the important aspects are shortly. Before we do so, we need to discuss a way to compare the one dimensional data vectors \mathbf{x}_1 and \mathbf{x}_2 .

Project Activity 24.2. To compare the two one dimensional data vectors, we need to consider variance and covariance.

- (a) With data it is useful to know how spread out the data is, something the average doesn't tell us. For example, the data sets $[1 \ 2 \ 3]^T$ and $[-2 \ 0 \ 8]^T$ both have averages of 2, but the data in $[-2 \ 0 \ 8]^T$ is more spread out. Variance provides one measure of how spread out a one-dimensional data set $\mathbf{x} = [x_1 \ x_2 \ x_3 \ \cdots \ x_n]^T$ is. Variance is defined as

$$\text{var}(\mathbf{x}) = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{\mathbf{x}})^2.$$

The variance provides a measure of how far from the average the data is spread.¹

Determine the variances of the two data vectors \mathbf{x}_1 and \mathbf{x}_2 . Which is more spread out?

- (b) In general, we will have more than one-dimensional data, as in our SAT data set. It will be helpful to have a way to compare one-dimensional data sets to try to capture the idea of variance for different data sets – how much the data in two different data sets varies from the mean with respect to each other. One such measure is covariance – essentially the average of all corresponding products of deviations from the means. We define the covariance of two data vectors $\mathbf{x} = [x_1 \ x_2 \ \cdots \ x_n]^T$ and $\mathbf{y} = [y_1 \ y_2 \ \cdots \ y_n]^T$ as

$$\text{cov}(\mathbf{x}, \mathbf{y}) = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{\mathbf{x}})(y_i - \bar{\mathbf{y}}).$$

Determine all covariances

$$\text{cov}(\mathbf{x}_1, \mathbf{x}_1), \text{ cov}(\mathbf{x}_1, \mathbf{x}_2), \text{ cov}(\mathbf{x}_2, \mathbf{x}_1), \text{ and } \text{cov}(\mathbf{x}_2, \mathbf{x}_2).$$

¹It might seem that we should divide by n instead of $n - 1$ in the variance, but it is generally accepted to do this for reasons we won't get into. Suffice it to say that if we are using a sample of the entire population, then dividing by $n - 1$ provides a variance whose square root is closer to the standard deviation than we would get if we divide by n . If we are calculating the variance of an entire population, then we would divide by n .



How are $\text{cov}(\mathbf{x}_1, \mathbf{x}_2)$ and $\text{cov}(\mathbf{x}_2, \mathbf{x}_1)$ related? How are $\text{cov}(\mathbf{x}_1, \mathbf{x}_1)$ and $\text{cov}(\mathbf{x}_2, \mathbf{x}_2)$ related to variances?

- (c) What is most important about covariance is its sign. Suppose $\mathbf{y} = [y_1 \ y_2 \ \dots \ y_n]^T$, $\mathbf{z} = [z_1 \ z_2 \ \dots \ z_n]^T$ and $\text{cov}(\mathbf{y}, \mathbf{z}) > 0$. Then if y_i is larger than y_j it is likely that z_i is also larger than z_j . For example, if \mathbf{y} is a vector that records a persons height from age 2 to 10 and \mathbf{z} records tha same person's weight in the same years, we might expect that when y_i increases so does z_i . Similarly, if $\text{cov}(\mathbf{y}, \mathbf{z}) < 0$, then as one data set increases, the other decreases. As an example, if \mathbf{y} records the number of hours a student spends playing video games each semester ad \mathbf{z} gives the students GPA for each semester, then we might expect that z_i decreases as y_i increases. When $\text{cov}(\mathbf{y}, \mathbf{z}) = 0$, then \mathbf{y} and \mathbf{z} are said to be uncorrelated or independent of each other.

For our example \mathbf{x}_1 and \mathbf{x}_2 , what does $\text{cov}(\mathbf{x}_1, \mathbf{x}_2)$ tell us about the relationship between \mathbf{x}_1 and \mathbf{x}_2 ? Why should we expect this from the context?

- (d) The covariance gives us information about the relationships between the attributes. So instead of working with the original data, we work with the covariance data. If we have m data vectors $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_m$ in \mathbb{R}^n , the *covariance matrix* $C = [c_{ij}]$ satisfies $c_{ij} = \text{cov}(\mathbf{x}_i, \mathbf{x}_j)$. Calculate the covariance matrix for our SAT data. Then explain why $C = \frac{1}{n-1} \mathbf{X} \mathbf{X}^T$.

Recall that the goal of PCA is to find a matrix P such that $P\mathbf{X} = \mathbf{Y}$ where P transforms the data set to a coordinate system in which the important aspects of the data are revealed. We are now in a position to discuss what that means.

An ideal view of our data would be one in which we can see the direction of greatest variance and one that minimizes redundancy. With redundant data the variables are not independent – that is, covariance is nonzero. So we would like the covariances to all be zero (or as close to zero as possible) to remove redundancy in our data. In other words, we want covariance for \mathbf{Y} to be diagonal.

Project Activity 24.3. Consider the covariance matrix $C = \begin{bmatrix} 1760.18 & 1967.62 \\ 1967.62 & 2319.29 \end{bmatrix}$. Find a matrix P whose columns are unit vectors that diagonalizes C . Use technology as appropriate.

The matrix P from Project Activity 24.3 has some especially important properties: the columns of P are unit vectors and, moreover, $P^{-1} = P^T$ (you should check this - such a matrix is called an *orthogonal* matrix). For our purposes, we want to diagonalize $\mathbf{X} \mathbf{X}^T$ with $P \mathbf{X} \mathbf{X}^T P^{-1}$, so the matrix P that serve our purposes is the one whose *rows* are the eigenvectors of $\mathbf{X} \mathbf{X}^T$. To understand why this matrix is the one we want, recall that we want to have $P\mathbf{X} = \mathbf{Y}$, and we want to diagonalize $\mathbf{X} \mathbf{X}^T$ to a diagonal covariance matrix $\mathbf{Y} \mathbf{Y}^T$. In this situation we will have (recalling that $P^{-1} = P^T$)

$$\frac{1}{n-1} \mathbf{Y} \mathbf{Y}^T = \frac{1}{n-1} (P\mathbf{X})(P\mathbf{X})^T = \frac{1}{n-1} P (\mathbf{X} \mathbf{X}^T) P^T = P (\mathbf{X} \mathbf{X}^T) P^{-1}.$$

So the matrix P that we want is exactly the one that diagonalizes $\mathbf{X} \mathbf{X}^T$.

Project Activity 24.4. There are two useful ways we can interpret the results of our work so far. The eigenvector of $\mathbf{X} \mathbf{X}^T$ that corresponds to the largest (also called the *dominant*) eigenvalue λ_1



is $[-0.66 \ -0.76]^\top$. A plot of the centered data along with the eigenspace E_{λ_1} of XX^\top spanned by $\mathbf{v}_1 = [-0.66 \ -0.76]^\top$ is shown at left in Figure 24.2. The eigenvector \mathbf{v}_1 is called the *first principal component* of X . Notice that this line E_{λ_1} indicates the direction of greatest variation in the data. In other words, when we project the data points onto E_{λ_1} , as shown at right in Figure 24.2, the variation of the resulting points is larger than it is for any other line. In other words, the data is most spread out in this direction.

- (a) There is another way we can interpret this result. If we drop a perpendicular from one of our data points to the space E_{λ_1} it creates a right triangle with sides of length a , b , and c as illustrated in the middle of Figure 24.2. Use this idea to explain why maximizing the variation also minimizes the sum of the squares of the distances from the data points to this line. As a result, we have projected our two-dimensional data onto the one-dimensional space that maximizes the variance of the data.

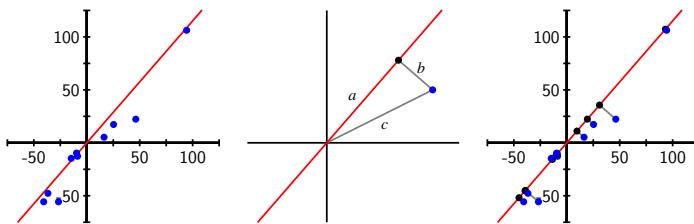


Figure 24.2: The principal component.

- (b) Recall that the matrix (to two decimal places) $P = \begin{bmatrix} -0.66 & -0.76 \\ -0.76 & 0.66 \end{bmatrix}$ transforms the data set X to a new data set Y whose covariance matrix is diagonal. Notice that P looks like a rotation matrix, but is not exactly a rotation matrix. To understand what P does to X , a plot of the original data from X is shown with solid blue circles and a plot of the transformed data from Y in solid magenta diamonds is at left in Figure 24.3. The form of the matrix P looks similar to a rotation matrix, but is not exactly a rotation matrix. Show that P is a combination of a rotation matrix (and find the rotation angle) and a reflection (and specifically identify the reflection). Explain how this relates to Figure 24.3. Also explain how the x -axis is related to the transformed data set Y .

The result of Project Activity 24.4 is that we have reduced the problem from considering the data in a two-dimensional space to a one-dimensional space E_{λ_1} where the most important aspect of the data is revealed. Of course, we eliminate some of the characteristics of the data, but the most important aspect is still included and highlighted.

The second eigenvector of XX^\top also has meaning. A picture of the eigenspace E_{λ_2} corresponding to the smaller eigenvector λ_2 of XX^\top is shown in Figure 24.4. The second eigenvector of XX^\top is perpendicular to the first, and the direction of the second eigenvector tells us the direction of the second most amount of variance as can be seen in Figure 24.4.

To summarize, the unit eigenvector for the largest eigenvalue of XX^\top indicates the direction in which the data has the greatest variance. The direction of the unit eigenvector for the smaller eigenvalue shows the direction in which the data has the second largest variance. This direction is

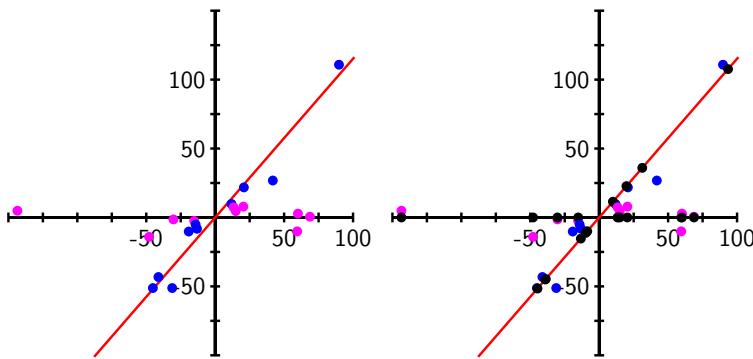
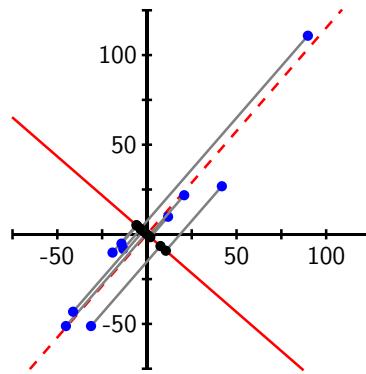
Figure 24.3: Applying P .

Figure 24.4: The second principal component.

also perpendicular to the first (indicating 0 covariance). The directions of the eigenvectors are called the *principal components* of X . The eigenvector with the highest eigenvalue is the first principal component of the data set and the other eigenvectors are ordered by the eigenvalues, highest to lowest. The principal components provide a new coordinate system from which to view our data – one in which we can see the maximum variance and in which there is zero covariance.

Project Activity 24.5. We can use the eigenvalues of XX^\top to quantify the amount of variance that is accounted for by our projections. Notice that the points along the x -axis at right in Figure 24.3 are exactly the numbers in the first row of Y . These numbers provide the projections of the data in Y onto the x -axis – the axis along which the data has its greatest variance.

- Calculate the variance of the data given by the first row of Y . This is the variance of the data in the direction of the eigenspace E_{λ_1} . How does the result compare to entries of the covariance matrix for Y .
- Repeat part (a) for the data along the second row of Y .
- The total variance of the data set is the sum of the variances. Calculate the sum of the variances for X and Y . What do you notice? Explain why the percentage of variance in

the data that is accounted for in the direction of E_{λ_1} is

$$\frac{\lambda_1}{\lambda_1 + \lambda_2}.$$

Then calculate this amount for the SAT data.

In general, PCA is most useful for larger data sets. The process is the same.

- Start with a set of data that forms the rows of an $m \times n$ matrix. We center the data by subtracting the mean of each row from the entries of that row to create a centered data set in a matrix X .
- The principal components of X are the eigenvectors of XX^T , ordered so that they correspond to the eigenvalues of XX^T in decreasing order.
- Let P be the matrix whose rows are the principal components of X , ordered from highest to lowest. Then $Y = PX$ is suitably transformed to identify the important aspects of the data.
- If $\lambda_1, \lambda_2, \dots, \lambda_n$ are the eigenvalues of XX^T in decreasing order, then the amount of variance in the data accounted for by the first r principal components is given by

$$\frac{\lambda_1 + \lambda_2 + \dots + \lambda_r}{\lambda_1 + \lambda_2 + \dots + \lambda_n}.$$

- The first r rows of $Y = PX$ provide the projection of the data set X onto an r -dimensional space spanned by the first r principal components of X .

Project Activity 24.6. Let us now consider a problem with more than two variables. We continue to keep the data set small so that we can conveniently operate with it. Table 24.2 presents additional information from ten states on four attributes related to the SAT – Participation rates, Evidence-Based Reading and Writing (EBRW) score, Math score, and average SAT score. Use technology as appropriate for this activity.

State	1	2	3	4	5	6	7	8	9	10
Rate	6	60	97	100	64	99	4	23	79	70
EBRW	595	540	522	508	565	512	643	574	534	539
Math	571	536	493	493	554	501	655	566	534	539
SAT	1166	1076	1014	1001	1120	1013	1296	1140	1068	1086

Table 24.2: SAT data.

- Determine the centered data matrix X for this data set.
- Find the covariance matrix for this data set. Round to four decimal places.
- Find the principal components of X . Include at least four decimal places accuracy.
- How much variation is accounted for in the data by the first principal component? In other words, if we reduce this data to one dimension, how much of the variation do we retain? Explain.



- (e) How much variation is accounted for in the data by the first two principal components? In other words, if we reduce this data to two dimensions, how much of the variation do we retain? Explain.



Section 25

Coordinate Vectors and Coordinate Transformations

Focus Questions

By the end of this section, you should be able to give precise and thorough answers to the questions listed below. You may want to keep these questions in mind to focus your thoughts as you complete the section.

- How do we find the coordinate vector of a vector \mathbf{x} with respect to a basis $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$?
- How can we visualize coordinate systems?
- How do we identify a vector space of dimension n with \mathbb{R}^n ?
- What properties does the coordinate transformation $\mathbf{x} \mapsto [\mathbf{x}]_{\mathcal{B}}$ have?
- How are the coordinates of a vector with respect to a basis related to its coordinates with respect to the standard basis?

Application: Calculating Sums

Consider the problem of calculating sums of powers of integers. For example,

$$\sum_{t=0}^{n-1} t = \frac{1}{2} (n^2 - n)$$

$$\sum_{t=0}^{n-1} t^2 = \frac{1}{6} (2n^3 - 3n^2 + 1)$$

$$\sum_{t=0}^{n-1} t^3 = \frac{1}{4} (n^4 - 2n^3 + n^2)$$

and so on. It is possible to prove these formulas if we have an idea of what the formula is, but how can one come up with these formulas in the first place? As we will see later in this section, we can make use of coordinate vectors to find and verify such formulas.

Introduction

In this section we will investigate how a basis of a vector space provides a coordinate system in which each vector in the vector space has a unique set of coordinates. Such a coordinate system will make any n -dimensional vector space look like \mathbb{R}^n . If the vector space is already \mathbb{R}^n , then each basis will provide us with a new perspective to visualize \mathbb{R}^n . We begin our analysis of coordinate systems by looking at how a basis in \mathbb{R}^2 gives us a different view of \mathbb{R}^2 .

Preview Activity 25.1. Two vectors \mathbf{v}_1 and \mathbf{v}_2 are shown in Figure 25.1.

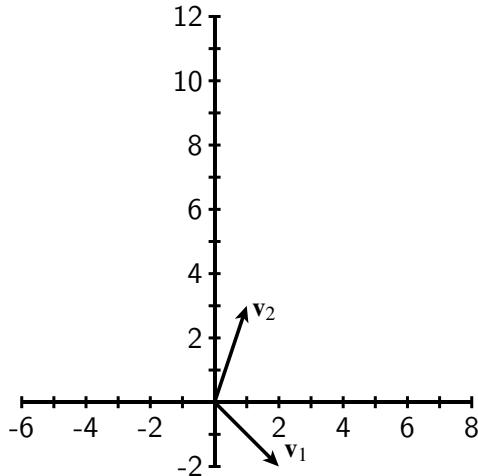


Figure 25.1: Vectors \mathbf{v}_1 and \mathbf{v}_2 .

- (1) Explain why $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2\}$ is a basis for \mathbb{R}^2 .
- (2) Draw the vector $\mathbf{b} = -3\mathbf{v}_1 + 2\mathbf{v}_2$ in Figure 25.1. Explain your process.
- (3) In part (1) we were given the weights $(-3 \text{ and } 2)$ of the linear combination of \mathbf{v}_1 and \mathbf{v}_2 that produced \mathbf{b} . We call the vector $\begin{bmatrix} -3 \\ 2 \end{bmatrix}$ the *coordinate vector* of \mathbf{b} with respect to the basis $\{\mathbf{v}_1, \mathbf{v}_2\}$. Explain why any vector \mathbf{b} in \mathbb{R}^2 can be written as a linear combination of vectors $\mathbf{v}_1, \mathbf{v}_2$. This shows that each vector in \mathbb{R}^2 has a coordinate vector in the coordinate system defined by \mathbf{v}_1 and \mathbf{v}_2 .
- (4) Since $\{\mathbf{v}_1, \mathbf{v}_2\}$ is a basis for \mathbb{R}^2 , any vector \mathbf{b} in \mathbb{R}^2 has a coordinate vector in the coordinate system defined by \mathbf{v}_1 and \mathbf{v}_2 . But we also need to make sure that each vector has a unique coordinate vector. Explain why there is no vector in \mathbb{R}^2 which has two different coordinate vectors.

- (5) We can think of the vectors \mathbf{v}_1 and \mathbf{v}_2 as defining a coordinate system, with $\text{Span}\{\mathbf{v}_1\}$ as the “ x ”-axis and $\text{Span}\{\mathbf{v}_2\}$ as the “ y ”-axis. Any vector \mathbf{b} in \mathbb{R}^2 can be written uniquely in the form

$$\mathbf{b} = x_1 \mathbf{v}_1 + x_2 \mathbf{v}_2$$

and the weights serve as the coordinates of \mathbf{b} in the \mathbf{v}_1 , \mathbf{v}_2 coordinate system. In this case the coordinate vector $\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ of \mathbf{b} with respect to the basis $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2\}$ is written as $[\mathbf{b}]_{\mathcal{B}}$.

Let $\mathcal{B} = \left\{ \begin{bmatrix} -1 \\ 4 \end{bmatrix}, \begin{bmatrix} 2 \\ 0 \end{bmatrix} \right\}$.

(a) Show that \mathcal{B} is a basis for \mathbb{R}^2 .

(b) Find $[\mathbf{b}]_{\mathcal{B}}$ if $\mathbf{b} = \begin{bmatrix} 0 \\ 6 \end{bmatrix}$. Draw a picture to illustrate how \mathbf{b} is expressed as a linear combination of the vectors in \mathcal{B} .

Bases as Coordinate Systems

Bases are useful for many reasons. A basis provides us with a unique representation of the elements in the space as linear combinations of the coordinate system defined by the basis vectors. The characterization of a vector space in terms of such a coordinate system will ultimately provide us with a powerful way to identify any vector space of dimension n geometrically and algebraically with \mathbb{R}^n through coordinate assignment.

As we saw in Preview Activity 25.1, we can think of a basis \mathcal{B} of \mathbb{R}^2 as determining a coordinate system of \mathbb{R}^2 . For example, let $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2\}$ where $\mathbf{v}_1 = \begin{bmatrix} 2 \\ -2 \end{bmatrix}$ and $\mathbf{v}_2 = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$. The vector $\mathbf{b} = \begin{bmatrix} -4 \\ 12 \end{bmatrix}$ can be written as $\mathbf{b} = -3\mathbf{v}_1 + 2\mathbf{v}_2$. Figure 25.2 shows that if we plot the point that is -3 units in the \mathbf{v}_1 direction (where a “unit” is a copy of \mathbf{v}_1) and 2 units in the \mathbf{v}_2 direction, then the result is the vector from the origin to point defined by \mathbf{b} .

As discussed in Preview Activity 25.1, we can think of the vectors \mathbf{v}_1 and \mathbf{v}_2 as defining a coordinate system, with $\text{Span}\{\mathbf{v}_1\}$ as the “ x ”-axis and $\text{Span}\{\mathbf{v}_2\}$ as the “ y ”-axis. Since \mathcal{B} is a basis, any vector \mathbf{b} in \mathbb{R}^2 can be written uniquely in the form

$$\mathbf{b} = x_1 \mathbf{v}_1 + x_2 \mathbf{v}_2$$

and the weights serve as the coordinates of \mathbf{b} in the \mathbf{v}_1 , \mathbf{v}_2 coordinate system. We call the vector $\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ the *coordinate vector* of \mathbf{b} with respect to the basis \mathcal{B} and write this vector as $[\mathbf{b}]_{\mathcal{B}}$.

This is actually a familiar idea, one we have used for years. The standard coordinates of a vector $\mathbf{a} = \begin{bmatrix} a_1 \\ a_2 \end{bmatrix}$ in \mathbb{R}^2 are just the coordinates of \mathbf{a} with respect to the standard basis $\{\mathbf{e}_1, \mathbf{e}_2\}$ of \mathbb{R}^2 .

While we can draw pictures in \mathbb{R}^2 , there is no reason to restrict this idea to \mathbb{R}^2 .



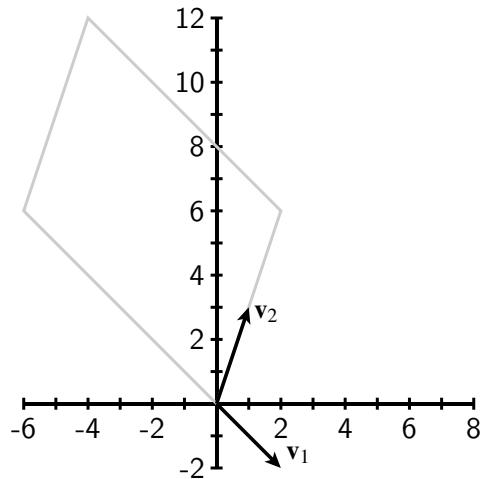


Figure 25.2: A Coordinate System Defined by a Basis.

Definition 25.1. Let $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$ be a basis for a vector space V . For any vector \mathbf{x} in V , the **coordinate vector** of \mathbf{x} with respect to \mathcal{B} is the vector

$$[\mathbf{x}]_{\mathcal{B}} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix},$$

where

$$\mathbf{x} = x_1 \mathbf{v}_1 + x_2 \mathbf{v}_2 + \cdots + x_n \mathbf{v}_n.$$

The scalars x_1, x_2, \dots, x_n are the **coordinates of the vector \mathbf{x} with respect to the basis**.

Recall that there is exactly one way to write a vector as a linear combination of basis vectors, so there is only one coordinate vector of a given vector with respect to a basis. Therefore, the coordinate vector of any vector is well-defined.

Activity 25.1. Let $\mathcal{S} = \{1, t\}$ and $\mathcal{B} = \{3 + 2t, 1 - t\}$. Assume that \mathcal{S} and \mathcal{B} are bases for \mathbb{P}_1 . Find $[3 + 7t]_{\mathcal{S}}$ and $[3 + 7t]_{\mathcal{B}}$. Note that the coordinate vector depends on the basis that is used.

IMPORTANT NOTE: We have defined the coordinate vector of a vector \mathbf{x} in a vector space V with respect to a basis $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$ as the vector $[x_1 \ x_2 \ \dots \ x_n]^T$ if

$$\mathbf{x} = x_1 \mathbf{v}_1 + x_2 \mathbf{v}_2 + \cdots + x_n \mathbf{v}_n.$$

Until now we have listed a basis as a set without regard to the order in which the basis elements are written. That is, the set $\{\mathbf{v}_1, \mathbf{v}_2\}$ is the same as the set $\{\mathbf{v}_2, \mathbf{v}_1\}$. Notice, however, that if we change the order of the vectors in our basis, say from $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$ to $\{\mathbf{v}_2, \mathbf{v}_1, \mathbf{v}_3, \dots, \mathbf{v}_n\}$, then the coordinate vector of \mathbf{x} with respect to \mathcal{B} will be different. To avoid this problem, when discussing coordinate vectors we will consider our bases to be *ordered bases*, so that the order in which we write the elements in our basis is fixed. So, for example, the ordered basis $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$ is different than the ordered basis $\{\mathbf{v}_2, \mathbf{v}_1, \mathbf{v}_3, \dots, \mathbf{v}_n\}$.

The Coordinate Transformation

We have seen that both \mathbb{R}^2 and \mathbb{P}_1 are vector spaces of dimension 2. Moreover, we can pair the polynomial $a_0 + a_1t$ in \mathbb{P}_1 with the vector $[a_0 \ a_1]$ in \mathbb{R}^2 as the coordinate vector of $a_0 + a_1t$ with respect to the standard basis for \mathbb{P}_1 . In this way we can see that the vector space \mathbb{P}_1 “looks” like the vector space \mathbb{R}^2 . In fact, using this same idea, we can show that any vector space of dimension n “looks” like \mathbb{R}^n . The key idea to make this work is the coordinate transformation.

As we saw in Activity 25.1, the coordinate vector of a vector with respect to a basis depends on the basis we choose. In this activity we found that $[3+7t]_{\mathcal{S}} = [3 \ 7]^T$ and $[3+7t]_{\mathcal{B}} = [2 \ -3]^T$, where $\mathcal{S} = \{1, t\}$ and $\mathcal{B} = \{3 + 2t, 1 - t\}$. These coordinate vectors allow us to identify polynomials in \mathbb{P}_1 with vectors in \mathbb{R}^2 . In this way, the coordinate vectors provide an identification between the vector space \mathbb{P}_1 and the vector space \mathbb{R}^2 via a *coordinate transformation*.

Definition 25.2. Let V be a vector space of dimension n with basis \mathcal{B} . The **coordinate transformation** T from V to \mathbb{R}^n with respect to the basis \mathcal{B} is the mapping defined by

$$T(\mathbf{x}) = [\mathbf{x}]_{\mathcal{B}}$$

for any vector \mathbf{x} in V .

Coordinate transformations have several useful properties that allow us to use them to compare the structure of vector spaces.

Activity 25.2. Let $a(t) = 8 + 2t$ and $b(t) = -5 + t$ in \mathbb{P}_1 . Suppose we know $\mathcal{C} = \{1 + t, 2 - t\}$ is a basis of \mathbb{P}_1 . Let $T(\mathbf{x}) = [\mathbf{x}]_{\mathcal{C}}$ for \mathbf{x} in \mathbb{P}_1 .

- (a) What are $T(a(t))$ and $T(b(t))$?
- (b) Find $T(a(t)) + T(b(t))$.
- (c) What is $T(a(t) + b(t)) = [a(t) + b(t)]_{\mathcal{C}}$?
- (d) What is the relationship between $T(a(t)) + T(b(t))$ and $T(a(t) + b(t))$?
- (e) Show that if c is any scalar, then $T(ca(t)) = cT(a(t))$.
- (f) Where have we seen functions with these properties before?

Activity 25.2 shows that coordinate transformations behave in ways similar to matrix transformations. Recall that if A is an $m \times n$ matrix, then

$$A(\mathbf{u} + \mathbf{w}) = A\mathbf{u} + A\mathbf{w} \text{ and } A(c\mathbf{u}) = cA\mathbf{u}$$

for any vectors \mathbf{u} and \mathbf{w} in \mathbb{R}^n and any scalar c . Because of these properties, the transformation A preserves linear combinations. That is, if $\mathbf{v}_1, \mathbf{v}_2, \dots$, and \mathbf{v}_n are any vectors in \mathbb{R}^n and c_1, c_2, \dots, c_n are any scalars, then

$$A(c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_n\mathbf{v}_n) = c_1A\mathbf{v}_1 + c_2A\mathbf{v}_2 + \cdots + c_nA\mathbf{v}_n.$$



Activity 25.2 contains the essential ideas to show that if T is a coordinate transformation from a vector space V with basis \mathcal{B} to \mathbb{R}^n , then

$$T(\mathbf{u} + \mathbf{w}) = T(\mathbf{u}) + T(\mathbf{w}) \text{ and } T(c\mathbf{u}) = cT(\mathbf{u})$$

for any vectors \mathbf{u} and \mathbf{w} in V and any scalar c . The fact that any coordinate transformation satisfies these properties is contained in the following theorem.

Theorem 25.3. *If a vector space V has a basis \mathcal{B} of n vectors, then the coordinate mapping $T : V \rightarrow \mathbb{R}^n$ defined by $T(\mathbf{x}) = [\mathbf{x}]_{\mathcal{B}}$ satisfies*

$$(1) \quad T(\mathbf{u} + \mathbf{w}) = T(\mathbf{u}) + T(\mathbf{w}) \text{ and}$$

$$(2) \quad T(c\mathbf{u}) = cT(\mathbf{u})$$

for any vectors \mathbf{u} and \mathbf{w} in V and any scalar c .

Proof. Let $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$ be vectors that form a basis \mathcal{B} for a vector space V . To verify the first property, let \mathbf{u} and \mathbf{w} be two arbitrary vectors from V . We will show that $T(\mathbf{u}) + T(\mathbf{w}) = T(\mathbf{u} + \mathbf{w})$ for these two vectors. Consider $T(\mathbf{u})$. If $T(\mathbf{u}) = [u_1 \ u_2 \ \dots \ u_n]^T$, then the fact that $T(\mathbf{u}) = [\mathbf{u}]_{\mathcal{B}}$ implies that

$$\mathbf{u} = u_1\mathbf{v}_1 + u_2\mathbf{v}_2 + \dots + u_n\mathbf{v}_n.$$

Similarly, if $T(\mathbf{w}) = [w_1 \ w_2 \ \dots \ w_n]^T$, then $T(\mathbf{w}) = [\mathbf{w}]_{\mathcal{B}}$ implies that

$$\mathbf{w} = w_1\mathbf{v}_1 + w_2\mathbf{v}_2 + \dots + w_n\mathbf{v}_n.$$

We then obtain

$$\begin{aligned} \mathbf{u} + \mathbf{w} &= (u_1\mathbf{v}_1 + u_2\mathbf{v}_2 + \dots + u_n\mathbf{v}_n) + (w_1\mathbf{v}_1 + w_2\mathbf{v}_2 + \dots + w_n\mathbf{v}_n) \\ &= (u_1 + w_1)\mathbf{v}_1 + (u_2 + w_2)\mathbf{v}_2 + \dots + (u_n + w_n)\mathbf{v}_n. \end{aligned}$$

Thus, by definition of T again,

$$T(\mathbf{u} + \mathbf{w}) = [\mathbf{u} + \mathbf{w}]_{\mathcal{B}} = [(u_1 + w_1) \ (u_2 + w_2) \ \dots \ (u_n + w_n)]^T.$$

To show that $T(\mathbf{u} + \mathbf{w}) = T(\mathbf{u}) + T(\mathbf{w})$, note that

$$\begin{aligned} T(\mathbf{u}) + T(\mathbf{w}) &= [u_1 \ u_2 \ \dots \ u_n]^T + [w_1 \ w_2 \ \dots \ w_n]^T \\ &= [(u_1 + w_1) \ (u_2 + w_2) \ \dots \ (u_n + w_n)]^T \\ &= [\mathbf{u} + \mathbf{w}]_{\mathcal{B}} \\ &= T(\mathbf{u} + \mathbf{w}). \end{aligned}$$

The proof of the second property is left for the exercises.



The first result of Theorem 25.3 can be extended to any finite number of vectors by mathematical induction. That is, if $T : V \rightarrow \mathbb{R}^n$ is the coordinate mapping defined by $T(\mathbf{x}) = [\mathbf{x}]_{\mathcal{B}}$ for some basis \mathcal{B} , and if $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_m$ are vectors in V , then $T(\mathbf{u}_1 + \mathbf{u}_2 + \dots + \mathbf{u}_m) = T(\mathbf{u}_1) + T(\mathbf{u}_2) + \dots + T(\mathbf{u}_m)$. In other words,

$$[\mathbf{u}_1 + \mathbf{u}_2 + \dots + \mathbf{u}_m]_{\mathcal{B}} = [\mathbf{u}_1]_{\mathcal{B}} + [\mathbf{u}_2]_{\mathcal{B}} + \dots + [\mathbf{u}_m]_{\mathcal{B}}$$

for any vectors $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_m$ in a vector space V with basis \mathcal{B} .

In essence, a coordinate transformation makes an identification of any vector space of dimension n with \mathbb{R}^n . One of the driving forces behind this identification is the fact that, just as in \mathbb{R}^n , if a vector space V has a basis of n elements, then any basis for V will have exactly n elements. The fact that any coordinate transformation from a vector space V to \mathbb{R}^n is linear means that elements in V behave the same way with respect to addition and multiplication by scalars in V as do their images in \mathbb{R}^n under the coordinate transformation. To make this identification complete, there are still two questions to address. First, given a coordinate transformation T from a vector space V to \mathbb{R}^n , is it possible for two vectors in V to have the same image in \mathbb{R}^n (in other words, is T one-to-one), and is every vector in \mathbb{R}^n the image of some vector in V under T (in other words, does T map V onto \mathbb{R}^n)? If the coordinate transformations are one-to-one and onto, then, in essence, any vector space of dimension n is just a copy of \mathbb{R}^n . As a result, to understand any vector space with a basis of n vectors, it is enough to understand \mathbb{R}^n .

Activity 25.3. Let V be a vector space with an ordered basis $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$. Then T maps V into \mathbb{R}^n . We want to show that T maps V onto \mathbb{R}^n . Recall that a function f from a set X to a set Y is onto if for any element y in Y , there is an element x in X such that $f(x) = y$. Let $\mathbf{b} = [b_1 \ b_2 \ \dots \ b_n]^T$ be a vector in \mathbb{R}^n . Must there be a vector \mathbf{v} in V so that $T(\mathbf{v}) = \mathbf{b}$? If so, find such a vector. If not, explain why not.

Activity 25.3 shows that a coordinate transformation maps an n -dimensional vector space V onto \mathbb{R}^n . What's more, any coordinate transformation is also one-to-one (the proof that T is one-to-one is left for the exercises). We summarize these results in the following theorem.

Theorem 25.4. *If a vector space V has a basis \mathcal{B} of n vectors, then the coordinate mapping $T : V \rightarrow \mathbb{R}^n$ defined by $T(\mathbf{x}) = [\mathbf{x}]_{\mathcal{B}}$ is both one-to-one and onto.*

Theorems 25.3 and 25.4 show that a coordinate transformation from an n -dimensional vector space V with basis \mathcal{B} to \mathbb{R}^n provides an identification of V with \mathbb{R}^n , where a vector \mathbf{v} in V is identified with the vector $[\mathbf{v}]_{\mathcal{B}}$. This identification allows us to transfer questions (linear dependence, independence, span) in V to \mathbb{R}^n where we can apply our knowledge of matrices to answer the questions.

Activity 25.4. Let $V = \mathbb{P}_3$ and let $\mathcal{B} = \{1, t, t^2, t^3\}$. Let $S = \{1 + t + t^2 + t^3, t - t^3, 1 + 2t^2, 1 + 5t - t^3\}$.

- (a) Find each of $[1 + t + t^2 + t^3]_{\mathcal{B}}$, $[t - t^3]_{\mathcal{B}}$, $[1 + 2t^2]_{\mathcal{B}}$, and $[1 + 5t - t^3]_{\mathcal{B}}$.
- (b) Are the vectors $[1 + t + t^2 + t^3]_{\mathcal{B}}$, $[t - t^3]_{\mathcal{B}}$, $[1 + 2t^2]_{\mathcal{B}}$, and $[1 + 5t - t^3]_{\mathcal{B}}$ linearly independent or dependent? Explain. If the vectors are linearly independent, write one of the vectors as a linear combination of the others.



- (c) The coordinate transformation identifies the vectors in $S = \{1 + t + t^2 + t^3, t - t^3, 1 + 2t^2, 1 + 5t - t^3\}$ with their coordinate vectors in \mathbb{R}^4 . Use that information to determine if S is a linearly independent or dependent set. If dependent, write one of the vectors in S as a linear combination of the others.

Examples

What follows are worked examples that use the concepts from this section.

Example 25.5.

- (a) Find the coordinate vector of \mathbf{v} with respect to the ordered basis \mathcal{B} in the indicated vector space.
- $\mathcal{B} = \{1 + t, 2 - t\}$ in \mathbb{P}_1 with $\mathbf{v} = 4 + t$
 - $\mathcal{B} = \left\{ \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \right\}$ in $\mathcal{M}_{2 \times 2}$ with $\mathbf{v} = \begin{bmatrix} 2 & 3 \\ 1 & 0 \end{bmatrix}$
- (b) Find the vector \mathbf{v} in the indicated vector space V given the basis \mathcal{B} of V and $[\mathbf{v}]_{\mathcal{B}}$.
- $V = \mathbb{P}_2$, $\mathcal{B} = \{1 + t^2, 1 + t, t + t^2\}$, $[\mathbf{v}]_{\mathcal{B}} = [2 \ 1 \ 3]^T$
 - $\mathcal{B} = \left\{ \cos(x), \frac{1}{1+x^2} \right\}$, $V = \text{Span } \mathcal{B}$ as a subspace of \mathcal{F} , $[\mathbf{x}]_{\mathcal{B}} = [2 \ -1]^T$

Example Solution.

- (a) Find the coordinate vector of \mathbf{v} with respect to the ordered basis \mathcal{B} in the indicated vector space.
- We need to write $\mathbf{v} = 4 + t$ as a linear combination of $1 + t$ and $2 - t$. If $4 + t = c_1(1+t) + c_2(2-t)$, then equating coefficients of like power terms yields the equations $4 = c_1 + 2c_2$ and $1 = c_1 - c_2$. The solution to this system is $c_1 = 2$ and $c_2 = 1$, so $[4 + t]_{\mathcal{B}} = [2 \ 1]^T$.
 - We need to write \mathbf{v} as a linear combination of the vectors in \mathcal{B} . If

$$\begin{bmatrix} 2 & 3 \\ 1 & 0 \end{bmatrix} = c_1 \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + c_2 \begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix} + c_3 \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} + c_4 \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix},$$

equating corresponding components produces the system

$$\begin{aligned} c_1 + c_2 + c_3 &= 2 \\ c_3 &= 3 \\ -c_2 &= 1 \\ c_1 + c_2 + c_3 + c_4 &= 0. \end{aligned}$$

The solution to this system is $c_1 = 0$, $c_2 = -1$, $c_3 = 3$, and $c_4 = -2$, so $[\mathbf{v}]_{\mathcal{B}} = [0 \ -1 \ 3 \ -2]^T$.



(b) Find the vector \mathbf{v} in the indicated vector space V given the basis \mathcal{B} of V and $[\mathbf{v}]_{\mathcal{B}}$.

i. Since $[\mathbf{v}]_{\mathcal{B}} = [2 \ 1 \ 3]^T$, it follows that

$$\mathbf{v} = 2(1 + t^2) + 1(1 + t) + 3(t + t^2) = 3 + 4t + 5t^2.$$

ii. Since $[\mathbf{x}]_{\mathcal{B}} = [2 \ -1]^T$, it follows that

$$\mathbf{v} = 2 \cos(x) - \frac{1}{1+x^2}.$$

Example 25.6. Let $p_1(t) = 1$, $p_2(t) = 2 - t$, $p_3(t) = 3 + t - t^2$, $p_4(t) = t + t^3$, and $p_5(t) = 2t + t^2 + t^4$ be in \mathbb{P}_4 . Also, let $\mathcal{B} = \{1, t, t^2, t^3, t^4\}$ be the standard basis for \mathbb{P}_4 .

- (a) Find $[p_1(t)]_{\mathcal{B}}, [p_2(t)]_{\mathcal{B}}, [p_3(t)]_{\mathcal{B}}, [p_4(t)]_{\mathcal{B}}$, and $[p_5(t)]_{\mathcal{B}}$.
- (b) Use the result of part (a) to explain why $\{p_1(t), p_2(t), p_3(t), p_4(t), p_5(t)\}$ is a basis for \mathbb{P}_4 .
- (c) Let $p(t) = 2 - t + t^2 - 3t^3 + 4t^4$. Find $[p(t)]_{\mathcal{B}}$.
- (d) Use the coordinate vectors in parts (a) and (c) to write $p(t)$ as a linear combination of $p_1(t), p_2(t), p_3(t), p_4(t)$, and $p_5(t)$

Example Solution.

- (a) The coordinate vectors of a polynomial with respect to the standard basis in \mathbb{P}_4 is found by just reading off the coefficients of the polynomial. So

$$\begin{aligned} [p_1(t)]_{\mathcal{B}} &= [1 \ 0 \ 0 \ 0 \ 0]^T & [p_2(t)]_{\mathcal{B}} &= [2 \ -1 \ 0 \ 0 \ 0]^T \\ [p_3(t)]_{\mathcal{B}} &= [3 \ 1 \ -1 \ 0 \ 1]^T & [p_4(t)]_{\mathcal{B}} &= [0 \ 1 \ 0 \ 1 \ 0]^T \\ [p_5(t)]_{\mathcal{B}} &= [0 \ 2 \ 1 \ 0 \ 1]^T. \end{aligned}$$

- (b) Let

$$\begin{aligned} \mathcal{C} &= \{p_1(t), p_2(t), p_3(t), p_4(t), p_5(t)\} \text{ and} \\ \mathcal{S} &= \{[p_1(t)]_{\mathcal{B}}, [p_2(t)]_{\mathcal{B}}, [p_3(t)]_{\mathcal{B}}, [p_4(t)]_{\mathcal{B}}, [p_5(t)]_{\mathcal{B}}\}. \end{aligned}$$

Since the coordinate transformation is one-to-one and onto, the two sets \mathcal{C} and \mathcal{S} are either both linearly dependent or linearly independent. Technology shows that the reduced row echelon form of

$$[[p_1(t)]_{\mathcal{B}} \ [p_2(t)]_{\mathcal{B}} \ [p_3(t)]_{\mathcal{B}} \ [p_4(t)]_{\mathcal{B}} \ [p_5(t)]_{\mathcal{B}}]$$

is I_5 , so the sets \mathcal{C} and \mathcal{S} are both linearly independent. Since $\dim(\mathbb{P}_4) = 5$, it follows that any linearly independent set of five vectors in \mathbb{P}_4 is a basis for \mathbb{P}_4 . Therefore, the set \mathcal{C} is a basis for \mathbb{P}_4 .

- (c) Similar to part (a), we have $[p(t)]_{\mathcal{B}} = [2 \ -1 \ 1 \ -3 \ 4]^T$.



(d) Technology shows that the reduced row echelon form of the augmented matrix

$$[[p_1(t)]_{\mathcal{B}} \ [p_2(t)]_{\mathcal{B}} \ [p_3(t)]_{\mathcal{B}} \ [p_4(t)]_{\mathcal{B}} \ [p_5(t)]_{\mathcal{B}} \ | \ [p(t)]_{\mathcal{B}}]$$

is

$$\left[\begin{array}{ccccc|c} 1 & 0 & 0 & 0 & 0 & -25 \\ 0 & 1 & 0 & 0 & 0 & 9 \\ 0 & 0 & 1 & 0 & 0 & 3 \\ 0 & 0 & 0 & 1 & 0 & -3 \\ 0 & 0 & 0 & 0 & 1 & 4 \end{array} \right].$$

So

$$[p(t)]_{\mathcal{B}} = -25[p_1(t)]_{\mathcal{B}} + 9[p_2(t)]_{\mathcal{B}} + 3[p_3(t)]_{\mathcal{B}} - 3[p_4(t)]_{\mathcal{B}} + 4[p_5(t)]_{\mathcal{B}}$$

and

$$p(t) = -25p_1(t) + 9p_2(t) + 3p_3(t) - 3p_4(t) + 4p_5(t).$$

Summary

The key idea in this handout is the coordinate vector with respect to a basis.

- If $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \dots, \mathbf{v}_n\}$ is a basis for a vector space V , then the coordinate vector of \mathbf{x} in V with respect to \mathcal{B} is the vector

$$[\mathbf{x}]_{\mathcal{B}} = [x_1, x_2, \dots, x_n]^T,$$

where

$$\mathbf{x} = x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + \cdots + x_n\mathbf{v}_n.$$

- The coordinate transformation $\mathbf{x} \mapsto [\mathbf{x}]_{\mathcal{B}}$ is a one-to-one and onto transformation from an n -dimensional vector space V to \mathbb{R}^n which preserves linear combinations.
- The coordinate transformation $\mathbf{x} \mapsto [\mathbf{x}]_{\mathcal{B}}$ allows us to translate problems in arbitrary vector spaces to \mathbb{R}^n where we have already developed tools to solve the problems.

Exercises

- (1) Let $\mathcal{B} = \left\{ \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix} \right\}$ be a basis of the subspace defined by the equation $y - 4x + z = 0$. Find the coordinates of the vector $\mathbf{b} = \begin{bmatrix} 3 \\ 4 \\ 2 \end{bmatrix}$ with respect to the basis \mathcal{B} .

$$\text{0. Find the coordinates of the vector } \mathbf{b} = \begin{bmatrix} 3 \\ 4 \\ 2 \end{bmatrix} \text{ with respect to the basis } \mathcal{B}.$$

- (2) Given basis $\mathcal{B} = \{1 + t, 2 + t^2, t + t^2\}$ of \mathbb{P}_2 ,

- (a) For which $p(t)$ in \mathbb{P}_2 is $[p(t)]_{\mathcal{B}} = [1 \ -1 \ 3]^T$?

- (b) Determine coordinates of $q(t) = -1 + t + 2t^2$ in \mathbb{P}_2 with respect to basis \mathcal{B} .



(3) Find two different bases \mathcal{B}_1 and \mathcal{B}_2 of \mathbb{R}^2 so that $[\mathbf{b}]_{\mathcal{B}_1} = [\mathbf{b}]_{\mathcal{B}_2} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$, where $\mathbf{b} = \begin{bmatrix} 5 \\ 3 \end{bmatrix}$.

(4) If $[\mathbf{b}_1]_{\mathcal{B}} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ and $[\mathbf{b}_2]_{\mathcal{B}} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$ with respect to some basis \mathcal{B} , where $\mathbf{b}_1 = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$ and $\mathbf{b}_2 = \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix}$, what are the coordinates of $\begin{bmatrix} -2 \\ 3 \\ 1 \end{bmatrix}$?

(5) If $[\mathbf{b}_1]_{\mathcal{B}} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ and $[\mathbf{b}_2]_{\mathcal{B}} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$ with respect to some basis \mathcal{B} , where $\mathbf{b}_1 = \begin{bmatrix} 3 \\ 1 \\ 3 \end{bmatrix}$ and $\mathbf{b}_2 = \begin{bmatrix} 4 \\ 1 \\ 5 \end{bmatrix}$, what are the vectors in \mathcal{B} ?

(6) Let $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$ be a basis for a vector space V . Describe how the coordinates of a vector with respect to \mathcal{B} will change if \mathbf{v}_1 is replaced with $\frac{1}{2}\mathbf{v}_1$.

(7) Let $\mathcal{B} = \{1, t, 1 + t^2\}$ in \mathbb{P}_2 .

(a) Show that \mathcal{B} is a basis for \mathbb{P}_2 .

(b) Let $p_1(t) = 1 + 2t^2$, $p_2(t) = 1 + t + 2t^2$, and $p_3(t) = 2 - t + t^2$ in \mathbb{P}_2 .

i. Find $[p_1(t)]_{\mathcal{B}}$, $[p_2(t)]_{\mathcal{B}}$, and $[p_3(t)]_{\mathcal{B}}$.

ii. Use the coordinate vectors in part i. to determine if the set $\{p_1(t), p_2(t), p_3(t)\}$ is linearly independent or dependent.

(8) Let $W = \text{Span}\{2 + 4t + 6t^3, 3 - t^2, 3 - t^2 + 9t^3\}$ in \mathbb{P}_3 . Let $\mathcal{B} = \{1, t, t^2, t^3\}$ be the standard basis for \mathbb{P}_3 .

(a) Calculate $[2 + 4t + 6t^3]_{\mathcal{B}}$, $[3 - t^2]_{\mathcal{B}}$ and $[3 - t^2 + 9t^3]_{\mathcal{B}}$.

(b) Use the coordinate vectors from part (a) to determine if the polynomials $2 + 4t + 6t^3$, $3 - t^2$, and $3 - t^2 + 9t^3$ are linearly independent or dependent.

(c) Let $p(t) = 4 + 2t - t^2 + 9t^3$. Find $[p(t)]_{\mathcal{B}}$.

(d) Use the calculations from parts (a) and (c) to determine if $p(t)$ is in W . If so, write $p(t)$ as a linear combination of the polynomials $2 + 4t + 6t^3$, $3 - t^2$, and $3 - t^2 + 9t^3$. If not, explain why not.

(9) Let $\mathcal{B} = \left\{ \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix} \right\}$ in $M_{2 \times 2}$.

(a) Show that \mathcal{B} is a basis of $M_{2 \times 2}$.

(b) Let

$$A = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}, B = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, C = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}, D = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}.$$

Find $[A]_{\mathcal{B}}$, $[B]_{\mathcal{B}}$, $[C]_{\mathcal{B}}$, $[D]_{\mathcal{B}}$.



- (c) Determine if the set $\{A, B, C, D\}$ is linearly dependent or independent.
- (10) Let V be a vector space of dimension n and let \mathcal{B} be a basis for V . Show that the coordinate transformation T from V to \mathbb{R}^n defined by $T(\mathbf{x}) = [\mathbf{x}]_{\mathcal{B}}$ satisfies $T(\mathbf{0}_V) = \mathbf{0}$, where $\mathbf{0}_V$ is the additive identity in V .
- (11) Prove the second property of Theorem 25.3. That is, if a vector space V has a basis of n vectors, then the coordinate mapping $T : V \rightarrow \mathbb{R}^n$ defined by $T(\mathbf{x}) = [\mathbf{x}]_{\mathcal{B}}$ satisfies
- $$T(c\mathbf{u}) = cT(\mathbf{u})$$
- for any vector \mathbf{u} in V and any scalar c .
- (12) Prove Theorem 25.4 by demonstrating that if V is a vector space with ordered basis $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$, then the coordinate mapping $T : V \rightarrow \mathbb{R}^n$ defined by $T(\mathbf{x}) = [\mathbf{x}]_{\mathcal{B}}$ is one-to-one.
- (13) The coordinate transformation T is one-to-one, so it has an inverse T^{-1} . Let V be an n -dimensional vector space that has a basis \mathcal{B} , and let $T : V \rightarrow \mathbb{R}^n$ be the coordinate transformation defined by $T(\mathbf{x}) = [\mathbf{x}]_{\mathcal{B}}$. Let $S = \{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k\}$ be a subset of V and let $R = \{[\mathbf{u}_1]_{\mathcal{B}}, [\mathbf{u}_2]_{\mathcal{B}}, \dots, [\mathbf{u}_k]_{\mathcal{B}}\}$ in \mathbb{R}^n .
- (a) Suppose \mathbf{x} is in V with $\mathbf{x} = x_1\mathbf{u}_1 + x_2\mathbf{u}_2 + \dots + x_k\mathbf{u}_k$. Write the vector $[\mathbf{x}]_{\mathcal{B}}$ as a linear combination of the vectors in R . Explain your reasoning and explain how your result shows that T preserves linear combinations.
 - (b) Now suppose \mathbf{w} is in V so that $[\mathbf{w}]_{\mathcal{B}} = w_1[\mathbf{u}_1]_{\mathcal{B}} + w_2[\mathbf{u}_2]_{\mathcal{B}} + \dots + w_k[\mathbf{u}_k]_{\mathcal{B}}$. Write the vector \mathbf{w} as a linear combination of the vectors in S . Explain your reasoning and explain how your result shows that T^{-1} preserves linear combinations.
- (14) Let $S = \{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k\}$ be a subset of an n -dimensional vector space V with basis \mathcal{B} . Let $R = \{[\mathbf{u}_1]_{\mathcal{B}}, [\mathbf{u}_2]_{\mathcal{B}}, \dots, [\mathbf{u}_k]_{\mathcal{B}}\}$ in \mathbb{R}^n .
- (a) Show that if S is linearly independent in V , then R is linearly independent in \mathbb{R}^n .
 - (b) Is the converse of part (a) true? That is, if R is linearly independent in \mathbb{R}^n , must S be linearly independent in V ? Justify your answer.
 - (c) Repeat parts (a) and (b), replacing ‘linearly independent’ with ‘linearly dependent’.
- (15) Let V be an n -dimensional vector space with basis \mathcal{B} , and let $S = \{\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_k\}$ be a subset of V that spans V . Prove that $\{[\mathbf{w}_1]_{\mathcal{B}}, [\mathbf{w}_2]_{\mathcal{B}}, \dots, [\mathbf{w}_k]_{\mathcal{B}}\}$ spans \mathbb{R}^n .
- (16) Suppose \mathcal{B}_1 is a basis of a vector space V with n elements and \mathcal{B}_2 is a basis of W with n elements. Show that the map T_{VW} which sends every \mathbf{x} in V to the vector \mathbf{y} in W such that $[\mathbf{x}]_{\mathcal{B}_1} = [\mathbf{y}]_{\mathcal{B}_2}$ is one-to-one and onto.
- (17) Label each of the following statements as True or False. Provide justification for your response.
- (a) **True/False** The coordinates of a non-zero vector cannot be the same in the coordinate systems defined by two different bases.



- (b) **True/False** The coordinate vector of the zero vector with respect to any basis is always the zero vector.
- (c) **True/False** If W is a k dimensional subset of an n dimensional vector space V , and \mathcal{B} is a basis of W , then $[\mathbf{w}]_{\mathcal{B}}$ is a vector in \mathbb{R}^n for any \mathbf{w} in W .
- (d) **True/False** The order of vectors in a basis do not affect the coordinates of vectors with respect to this basis.
- (e) **True/False** If T is a coordinate transformation from a vector space V with basis \mathcal{B} to \mathbb{R}^n , then the vector $[\mathbf{x}]_{\mathcal{B}}$ is unique to the vector \mathbf{x} in V .
- (f) **True/False** If T is a coordinate transformation from a vector space V with basis \mathcal{B} to \mathbb{R}^n , then there is always a vector \mathbf{x} in V that maps to any vector \mathbf{b} in \mathbb{R}^n .
- (g) **True/False** A coordinate transformation from a vector space V with basis \mathcal{B} to \mathbb{R}^n always maps the additive inverse of a vector \mathbf{x} in V to the additive inverse of the vector $[\mathbf{x}]_{\mathcal{B}}$ in \mathbb{R}^n .
- (h) **True/False** A coordinate transformation provides a unique identification of vectors in an n -dimensional vector space with vectors in \mathbb{R}^n in a way that preserves the algebraic structure of the spaces.
- (i) **True/False** If the coordinate vector of \mathbf{x} in a vector space V is $\begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix}$, then the coordinate vector of $2\mathbf{x}$ is $\begin{bmatrix} 2 \\ -2 \\ 4 \end{bmatrix}$.

Project: Finding Formulas for Sums of Powers

One way to derive formulas for sums of powers of whole numbers is to use different bases and coordinate vectors. One basis that will be useful is a basis of polynomials created by the binomial coefficients. Recall that the binomial coefficient $\binom{n}{k}$ is equal to

$$\binom{n}{k} = \frac{n!}{k!(n-k)!},$$

with $\binom{n}{k} = 0$ if $n < k$.

The binomial coefficient can be rewritten in a way to make it applicable to polynomials as

$$\binom{n}{k} = \frac{n(n-1)(n-2)\cdots(n-k+1)}{k!} = \prod_{i=1}^n \frac{1}{i}(n-i+1).$$

With this representation of $\binom{n}{k}$, we can define a new polynomial in t of degree k as

$$p_k(t) = \frac{1}{k!}(t)(t-1)(t-2)\cdots(t-k+1)$$



with $p_0(t) = 1$. For example,

$$\begin{aligned} p_1(t) &= \frac{1}{1!}t = t \\ p_2(t) &= \frac{1}{2}(t)(t-1) = \frac{1}{2}(t^2 - t) \\ p_3(t) &= \frac{1}{6}(t)(t-1)(t-2) = \frac{1}{6}(t^3 - 3t^2 + 2t). \end{aligned}$$

These “generalized binomial coefficients” appear in Newton’s generalized binomial theorem.

Two facts will make these polynomials useful for our sums.

Project Activity 25.1. Our polynomials $p_k(t)$ are defined in terms of binomial coefficients. A useful identity will relate sums of binomial coefficients to other binomial coefficients. This identity, called the *hockey-stick identity* after the way it can be visualized on Pascal’s triangle, is as follows:

$$\sum_{k=0}^{n-1} \binom{k}{r} = \binom{n}{r+1}$$

for positive integers r . Use the definition of the binomial coefficients and some algebra to verify the hockey-stick identity.

The second useful fact about our polynomials $p_k(t)$ is that they form a basis for \mathbb{P}_n .

Project Activity 25.2.

- (a) Let k be a positive integer. Explain why $p_k(0) = p_k(1) = p_k(2) = \dots = p_k(k-1) = 0$ and $p_k(k) = 1$.
- (b) Let $\mathcal{P}_n = \{p_0(t), p_1(t), p_2(t), \dots, p_n(t)\}$. Show that \mathcal{P}_n is a basis for \mathbb{P}_n . (Hint: Let c_1, c_2, \dots, c_n be scalars and consider the equation

$$c_0p_0(t) + c_1p_1(t) + c_2p_2(t) + \dots + c_np_n(t) = 0.$$

Evaluate this equation at $t = 0, t = 1, \dots, t = n$ and use the result of part (a).)

Now we have the tools we need to derive our formulas. To simplify computations, we will change coordinates to the standard basis $\mathcal{S}_n = \{1, t, t^2, \dots, t^n\}$ for \mathbb{P}_n .

We will illustrate the process of deriving formulas for our sums with the sum $\sum_{t=0}^{n-1} t$. We want to write t as a linear combination of the vectors in \mathcal{P}_1 so that we can utilize the hockey-stick identity. In this case, by definition we have $t = p_1(t)$. It follows by the hockey-stick identity and the fact that $p_1(t) = \binom{t}{1}$ that

$$\begin{aligned} \sum_{t=0}^{n-1} t &= \sum_{t=0}^{n-1} p_1(t) \\ &= \sum_{t=0}^{n-1} \binom{t}{1} \\ &= \binom{n}{2} \\ &= \frac{1}{2}(n^2 - n). \end{aligned}$$



This is exactly the formula we saw at the beginning of this section. The cases for sums of higher powers work the same way, but we will need to do a little more work to write t^n in terms of the basis vectors in \mathcal{P}_n .

Project Activity 25.3. Consider the sum $\sum_{t=0}^{n-1} t^2$. We want to write t^2 as a linear combination of $p_0(t)$, $p_1(t)$ and $p_2(t)$. To do so, we will use the coordinate vectors with respect to \mathcal{S}_2 and do our work in \mathbb{R}^3 .

- (a) Find $[p_0(t)]_{\mathcal{S}_2}$, $[p_1(t)]_{\mathcal{S}_2}$, $[p_2(t)]_{\mathcal{S}_2}$, and $[t^2]_{\mathcal{S}_2}$.
- (b) Use the coordinate vectors from part (a) to write t^2 as a linear combination of the vectors in \mathcal{P}_2 .
- (c) Use the result of part (b) and, the hockey-stick identity, and the fact that $p_k(t) = \binom{t}{k}$ to find a formula for $\sum_{t=0}^{n-1} t^2$.

Deriving formulas for higher powers involves the same process, just with more algebra.

Project Activity 25.4. Use the process outlined in Project Activity 25.3 to derive formulas for the following sums.

- (a) $\sum_{t=0}^{n-1} t^3$
- (b) $\sum_{t=0}^{n-1} t^4$

Section 26

Change of Basis

Focus Questions

By the end of this section, you should be able to give precise and thorough answers to the questions listed below. You may want to keep these questions in mind to focus your thoughts as you complete the section.

- What is a change of basis matrix?
- Why is a change of basis useful?
- How can we use coordinate transformations to find a change of basis matrix?
- What are three important properties of change of basis matrices?

Application: Describing Orbits of Planets

Consider a planet orbiting the sun (or an object like a satellite orbiting the Earth). According to Kepler's Laws, we assume an elliptical orbit. There are many different ways to describe this orbit, and which description we use depends on our perspective and the application. One important perspective is to make the description of the orbit as simple as possible for earth-based observations. Two problems arise. One is that the earth's orbit and the orbit of the planet do not lie in the same plane. A second problem is that it is complicated to describe the orbit of a planet using the perspective of the plane of the earth's orbit. A reasonable approach, then, is to establish two different coordinate systems, one for the earth's orbit and one for the planet's orbit. We can then use a change of basis to move back and forth from these two perspectives.

Introduction

In calculus we changed coordinates, from rectangular to polar, for example, to make certain calculations easier. In order for us to be able to work effectively in different coordinate systems, and to

easily change back and forth as needed, we will want to have a way to effectively transition from one coordinate system to another. In other words, if we have two different bases for a vector space V , we want a straightforward way to translate between the coordinate vectors of any given vector in V with respect to the two bases.

Preview Activity 26.1.

- (1) Let $b_1(t) = 4 + 2t$, $b_2(t) = -6 + 8t$, $c_1(t) = 1 + t$, $c_2(t) = 1 - t$, and let $\mathcal{B} = \{b_1(t), b_2(t)\}$ and $\mathcal{C} = \{c_1(t), c_2(t)\}$.
- Show that \mathcal{B} and \mathcal{C} are bases for \mathbb{P}_1 .
 - Let $p(t) = 3b_1(t) + 2b_2(t)$. What is $[p(t)]_{\mathcal{B}}$?
 - Since \mathcal{C} is also a basis for \mathbb{P}_1 , there is also a coordinate vector for $p(t)$ with respect to \mathcal{C} , and it is reasonable to ask how $[p(t)]_{\mathcal{C}}$ is related to $[p(t)]_{\mathcal{B}}$. Recall that a coordinate transformation respects linear combinations – that is

$$[r\mathbf{x} + s\mathbf{y}]_{\mathcal{S}} = r[\mathbf{x}]_{\mathcal{S}} + s[\mathbf{y}]_{\mathcal{S}}$$

for any vectors \mathbf{x} and \mathbf{y} in a vector space with basis \mathcal{S} , and any scalars r and s . Use the fact that $p(t) = 3b_1(t) + 2b_2(t)$ and the linearity of the coordinate transformation with respect to the basis \mathcal{C} to express $[p(t)]_{\mathcal{C}}$ in terms of $[b_1(t)]_{\mathcal{C}}$ and $[b_2(t)]_{\mathcal{C}}$ (don't actually calculate $[b_1(t)]_{\mathcal{C}}$ and $[b_2(t)]_{\mathcal{C}}$ yet, just leave your result in terms of the symbols $[b_1(t)]_{\mathcal{C}}$ and $[b_2(t)]_{\mathcal{C}}$.)

- The result of part (c) can be expressed as a matrix-vector product of the form

$$[p(t)]_{\mathcal{C}} = P[p(t)]_{\mathcal{B}}.$$

Describe how the columns of the matrix P are related to $[b_1(t)]_{\mathcal{C}}$ and $[b_2(t)]_{\mathcal{C}}$.

- Now calculate $[b_1(t)]_{\mathcal{C}}$, $[b_2(t)]_{\mathcal{C}}$, and $[p(t)]_{\mathcal{C}}$. Determine the entries of the matrix P and verify in this example that $[p(t)]_{\mathcal{C}} = P[p(t)]_{\mathcal{B}}$.
- The matrix P that we constructed in problem (1) allows us to quickly and easily switch from coordinates with respect to a basis \mathcal{B} to coordinates with respect to another basis \mathcal{C} , providing a way to effectively transition from one coordinate system to another as described in the introduction. This matrix P is called a *change of basis matrix*. In problem (1) we explained why the change of basis matrix exists, and in this problem we will see another perspective from which to view this matrix. Let $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2\}$ and $\mathcal{C} = \{\mathbf{c}_1, \mathbf{c}_2\}$ be two bases for a vector space V . The change of basis matrix P from \mathcal{B} to \mathcal{C} has the property that $P[\mathbf{x}]_{\mathcal{B}} = [\mathbf{x}]_{\mathcal{C}}$ for every vector \mathbf{x} in V . We can determine the entries of P by applying this formula to specific vectors in V .
 - What are $[\mathbf{b}_1]_{\mathcal{B}}$ and $[\mathbf{b}_2]_{\mathcal{B}}$? Why?
 - If A is an $n \times n$ matrix and $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n$ are the standard unit vectors in \mathbb{R}^n (that is, \mathbf{e}_i is the i th column of the $n \times n$ identity matrix), then what does the product $A\mathbf{e}_i$ tell us about the matrix A ?
 - Combine the results of parts (a) and (b) and the equation $P[\mathbf{x}]_{\mathcal{B}} = [\mathbf{x}]_{\mathcal{C}}$ to explain why $P = [[\mathbf{b}_1]_{\mathcal{C}} \; [\mathbf{b}_2]_{\mathcal{C}}]$.



The Change of Basis Matrix

Suppose we have two different finite bases \mathcal{B} and \mathcal{C} for a vector space V . In Preview Activity 26.1 we learned how to translate between the two bases in the 2-dimensional case – if $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2\}$ and $\mathcal{C} = \{\mathbf{c}_1, \mathbf{c}_2\}$, then the change of basis matrix from \mathcal{B} to \mathcal{C} is the matrix $[[\mathbf{b}_1]_{\mathcal{C}} \ [\mathbf{b}_2]_{\mathcal{C}}]$. This result in the 2-dimensional case generalizes to the n -dimensional case, and we can determine a straightforward method for calculating a change of basis matrix. The essential idea was introduced in Preview Activity 26.1.

Let $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_n\}$ and $\mathcal{C} = \{\mathbf{c}_1, \mathbf{c}_2, \dots, \mathbf{c}_n\}$ be two bases for a vector space V . If \mathbf{x} is in V , we have defined the coordinate vectors $[\mathbf{x}]_{\mathcal{B}}$ and $[\mathbf{x}]_{\mathcal{C}}$ for \mathbf{x} with respect to \mathcal{B} and \mathcal{C} ,

respectively. Recall that $[\mathbf{x}]_{\mathcal{B}} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$ if

$$\mathbf{x} = x_1 \mathbf{b}_1 + x_2 \mathbf{b}_2 + \cdots + x_n \mathbf{b}_n.$$

To see how to convert from the coordinates of \mathbf{x} with respect to \mathcal{B} to coordinates of \mathbf{x} with respect to \mathcal{C} , note that

$$\begin{aligned} [\mathbf{x}]_{\mathcal{C}} &= [x_1 \mathbf{b}_1 + x_2 \mathbf{b}_2 + \cdots + x_n \mathbf{b}_n]_{\mathcal{C}} \\ &= x_1 [\mathbf{b}_1]_{\mathcal{C}} + x_2 [\mathbf{b}_2]_{\mathcal{C}} + \cdots + x_n [\mathbf{b}_n]_{\mathcal{C}} \\ &= [[\mathbf{b}_1]_{\mathcal{C}} \ [\mathbf{b}_2]_{\mathcal{C}} \ \cdots \ [\mathbf{b}_n]_{\mathcal{C}}] \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \\ &= [[\mathbf{b}_1]_{\mathcal{C}} \ [\mathbf{b}_2]_{\mathcal{C}} \ \cdots \ [\mathbf{b}_n]_{\mathcal{C}}] [\mathbf{x}]_{\mathcal{B}}. \end{aligned}$$

So we can convert from coordinates with respect to the basis \mathcal{B} to coordinates with respect to the basis \mathcal{C} by multiplying $[\mathbf{x}]_{\mathcal{B}}$ on the left by the matrix

$$[[\mathbf{b}_1]_{\mathcal{C}} \ [\mathbf{b}_2]_{\mathcal{C}} \ \cdots \ [\mathbf{b}_n]_{\mathcal{C}}].$$

This matrix is called the *change of basis matrix* from \mathcal{B} to \mathcal{C} and is denoted $P_{\mathcal{C} \leftarrow \mathcal{B}}$.

Definition 26.1. Let $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_n\}$ and $\mathcal{C} = \{\mathbf{c}_1, \mathbf{c}_2, \dots, \mathbf{c}_n\}$ be two bases for a vector space V . The **change of basis matrix** from \mathcal{B} to \mathcal{C} is the matrix

$$P_{\mathcal{C} \leftarrow \mathcal{B}} = [[\mathbf{b}_1]_{\mathcal{C}} \ [\mathbf{b}_2]_{\mathcal{C}} \ \cdots \ [\mathbf{b}_n]_{\mathcal{C}}].$$

The change of basis matrix allows us to convert from coordinates with respect to one basis to coordinates with respect to another. The result is summarized in the following theorem.

Theorem 26.2. Let $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_n\}$ and $\mathcal{C} = \{\mathbf{c}_1, \mathbf{c}_2, \dots, \mathbf{c}_n\}$ be two bases for a vector space V . Then

$$[\mathbf{x}]_{\mathcal{C}} = P_{\mathcal{C} \leftarrow \mathcal{B}} [\mathbf{x}]_{\mathcal{B}}$$

for any vector \mathbf{x} in V .



Finding a Change of Basis Matrix

We can use matrix techniques we have developed to quickly find a change of basis matrix in a finite dimensional vector space. The coordinate transformation will allow us to transfer our calculations into \mathbb{R}^n , and especially useful are the coordinate vectors with respect to a standard basis. These coordinate vectors are easy to find – for example, if $p(t) = 1 + t - 2t^2$ in \mathbb{P}_2 and $\mathcal{S} = \{1, t, t^2\}$ is the standard basis for \mathbb{P}_2 , then we can see by inspection that $[p(t)]_{\mathcal{S}} = [1 \ 1 \ -2]^T$. We will also utilize the same process we developed for calculating the inverse of an invertible $n \times n$ matrix A . Recall that this process involved solving the n matrix equations $Ax = \mathbf{e}_i$, where \mathbf{e}_i is the i th column of the $n \times n$ identity matrix I_n . We were able to solve these n systems all at one time by augmenting A with all of the vectors \mathbf{e}_i at once, which amounted to augmenting A with I_n and row reducing. Keep these ideas in mind for the following activity.

Activity 26.1. Let $b_1(t) = 4 + t$, $b_2(t) = 2 + 5t$, $c_1(t) = -1 + 2t$, and $c_2(t) = -1 - t$. The sets $\mathcal{B} = \{b_1(t), b_2(t)\}$ and $\mathcal{C} = \{c_1(t), c_2(t)\}$ are bases for \mathbb{P}_1 . Our goal is to find the change of basis matrix $P_{\mathcal{C} \leftarrow \mathcal{B}}$ from \mathcal{B} to \mathcal{C} . We will use the coordinate transformation with respect to the standard basis $\mathcal{S} = \{1, t\}$ for \mathbb{P}_1 to transfer our work into \mathbb{R}^2 .

- What are $[b_1(t)]_{\mathcal{S}}$, $[b_2(t)]_{\mathcal{S}}$, $[c_1(t)]_{\mathcal{S}}$, and $[c_2(t)]_{\mathcal{S}}$?
- What matrix equation must we solve to write $[b_1(t)]_{\mathcal{S}}$ as a linear combination of the vectors $[c_1(t)]_{\mathcal{S}}$ and $[c_2(t)]_{\mathcal{S}}$? How do we solve this equation? (Don't solve the equation yet.)
- What matrix equation must we solve to write $[b_2(t)]_{\mathcal{S}}$ as a linear combination of the vectors $[c_1(t)]_{\mathcal{S}}$ and $[c_2(t)]_{\mathcal{S}}$? How do we solve this equation? (Don't solve the equation yet.)
- Let A be the coefficient matrix of the systems you wrote in parts (b) and (c). Find the reduced row echelon form of the augmented matrix $[A \ | \ [b_1(t)]_{\mathcal{S}} \ [b_2(t)]_{\mathcal{S}}]$.
- Use the result of part (d) to write $b_1(t)$ and $b_2(t)$ as linear combinations of $c_1(t)$ and $c_2(t)$.
- Based on your responses to (b) and (c), if $[I_2 \ | \ P]$ is the reduced row echelon form of the matrix $[A \ | \ [b_1(t)]_{\mathcal{S}} \ [b_2(t)]_{\mathcal{S}}]$, what property will the matrix P have? Explain.

Activity 26.1 demonstrates how we can find a change of basis matrix in an n -dimensional vector space V . If $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_n\}$ and $\mathcal{C} = \{\mathbf{c}_1, \mathbf{c}_2, \dots, \mathbf{c}_n\}$ are two bases for V , then the change of basis matrix from \mathcal{B} to \mathcal{C} is given by

$$P_{\mathcal{C} \leftarrow \mathcal{B}} = [[\mathbf{b}_1]_{\mathcal{C}} \ [\mathbf{b}_2]_{\mathcal{C}} \ \cdots \ [\mathbf{b}_n]_{\mathcal{C}}].$$

To find $[\mathbf{b}_i]_{\mathcal{C}}$, we need to write \mathbf{b}_i as a linear combination of the vectors in \mathcal{C} . That is, we need to find weights $x_{1,i}, x_{2,i}, \dots, x_{n,i}$ so that

$$\mathbf{b}_i = x_{1,i}\mathbf{c}_1 + x_{2,i}\mathbf{c}_2 + \cdots + x_{n,i}\mathbf{c}_n. \quad (26.1)$$

The weights in equation (26.1) are also the weights that satisfy the equation

$$[\mathbf{b}_i]_{\mathcal{S}} = x_{1,i}[\mathbf{c}_1]_{\mathcal{S}} + x_{2,i}[\mathbf{c}_2]_{\mathcal{S}} + \cdots + x_{n,i}[\mathbf{c}_n]_{\mathcal{S}}$$



where \mathcal{S} is any basis for V . So to find these weights, we choose a convenient basis \mathcal{S} (often the standard basis, if one exists, is a good choice) and then row reduce the matrix

$$[[\mathbf{c}_1]_{\mathcal{S}} \ [\mathbf{c}_2]_{\mathcal{S}} \ \cdots \ [\mathbf{c}_n]_{\mathcal{S}} \mid [\mathbf{b}_i]_{\mathcal{S}}].$$

The row operations we will apply to row reduce the coefficient matrix

$$[[\mathbf{c}_1]_{\mathcal{S}} \ [\mathbf{c}_2]_{\mathcal{S}} \ \cdots \ [\mathbf{c}_n]_{\mathcal{S}}]$$

will be the same regardless of the augmented column, so we can solve all of the systems at one time by row reducing the matrix

$$[[\mathbf{c}_1]_{\mathcal{S}} \ [\mathbf{c}_2]_{\mathcal{S}} \ \cdots \ [\mathbf{c}_n]_{\mathcal{S}} \mid [\mathbf{b}_1]_{\mathcal{S}} \ [\mathbf{b}_2]_{\mathcal{S}} \ \cdots \ [\mathbf{b}_n]_{\mathcal{S}}].$$

The result of the row reduction will be the matrix

$$\left[I_n \mid \underset{\mathcal{C} \leftarrow \mathcal{B}}{P} \right].$$

Let us return to our example in Activity 26.1. For ease of computation we choose $\mathcal{S} = \{1, t\}$ to be the standard basis for \mathbb{P}_1 . Then $[\mathbf{b}_1(t)]_{\mathcal{S}} = [4 \ 1]^T$, $[\mathbf{b}_2(t)]_{\mathcal{S}} = [2 \ 5]^T$, $[\mathbf{c}_1(t)]_{\mathcal{S}} = [-1 \ 2]^T$, and $[\mathbf{c}_2(t)]_{\mathcal{S}} = [-1 \ -1]^T$. Row reducing

$$[[\mathbf{c}_1(t)]_{\mathcal{S}} \ [\mathbf{c}_2(t)]_{\mathcal{S}} \mid [\mathbf{b}_1(t)]_{\mathcal{S}} \ [\mathbf{b}_2(t)]_{\mathcal{S}}] = \left[\begin{array}{cc|cc} -1 & -1 & 4 & 2 \\ 2 & -1 & 1 & 5 \end{array} \right]$$

gives us

$$\left[\begin{array}{cc|cc} 1 & 0 & -1 & 1 \\ 0 & 1 & -3 & -3 \end{array} \right].$$

So the change of basis matrix from \mathcal{B} to \mathcal{C} is

$$\underset{\mathcal{C} \leftarrow \mathcal{B}}{P} = \left[\begin{array}{cc} -1 & 1 \\ -3 & -3 \end{array} \right].$$

Note that columns of $\underset{\mathcal{C} \leftarrow \mathcal{B}}{P}$ tell us how to write our basis vectors in \mathcal{B} as linear combinations of the basis vectors in \mathcal{C} . So we can check our work by noting that $b_1(t) = -c_1(t) - 3c_2(t)$ and $b_2(t) = c_1(t) - 3c_2(t)$.

The coordinate transformations, along with the change of basis matrix, allow us to visualize finite dimensional vector spaces in \mathbb{R}^n . In our example from Activity 26.1 we can view the vectors in the bases \mathcal{B} and \mathcal{C} as their coordinate vectors $[\mathbf{b}_1(t)]_{\mathcal{S}}$, $[\mathbf{b}_2(t)]_{\mathcal{S}}$, $[\mathbf{c}_1(t)]_{\mathcal{S}}$, and $[\mathbf{c}_2(t)]_{\mathcal{S}}$ as illustrated at left in Figure 26.1. The change of basis matrix $\underset{\mathcal{C} \leftarrow \mathcal{B}}{P}$ tells us how to transform a vector whose coordinates are given in the coordinate system defined by the basis \mathcal{B} to the coordinate system defined by the basis \mathcal{C} . For example, let $p(t) = 6 - 3t$. Then $[p(t)]_{\mathcal{S}} = [6 \ -3]^T$. We can find $[p(t)]_{\mathcal{B}}$ by row reducing the matrix $[[\mathbf{b}_1(t)]_{\mathcal{S}} \ [\mathbf{b}_2(t)]_{\mathcal{S}} \mid [p(t)]_{\mathcal{S}}]$ to its reduced row echelon form $\left[\begin{array}{cc|c} 1 & 0 & 2 \\ 0 & 1 & -1 \end{array} \right]$. So $[p(t)]_{\mathcal{S}} = 2[\mathbf{b}_1(t)]_{\mathcal{S}} - [\mathbf{b}_2(t)]_{\mathcal{S}}$ (and $p(t) = 2b_1(t) - b_2(t)$). Therefore, $[p(t)]_{\mathcal{B}} = [2 \ -1]^T$ as shown at right in Figure 26.1. Using the change of basis matrix we can see that

$$[p(t)]_{\mathcal{C}} = \underset{\mathcal{C} \leftarrow \mathcal{B}}{P} [p(t)]_{\mathcal{B}} = \left[\begin{array}{cc} -1 & 1 \\ -3 & -3 \end{array} \right] \left[\begin{array}{c} 2 \\ -1 \end{array} \right] = \left[\begin{array}{c} -3 \\ -3 \end{array} \right].$$

So $[p(t)]_{\mathcal{C}} = (-3)[\mathbf{c}_1(t)]_{\mathcal{S}} - 3[\mathbf{c}_2(t)]_{\mathcal{S}}$ (and $p(t) = (-3)c_1(t) - 3c_2(t)$). Therefore, $[p(t)]_{\mathcal{C}} = [-3 \ -3]^T$ as shown at right in Figure 26.1.



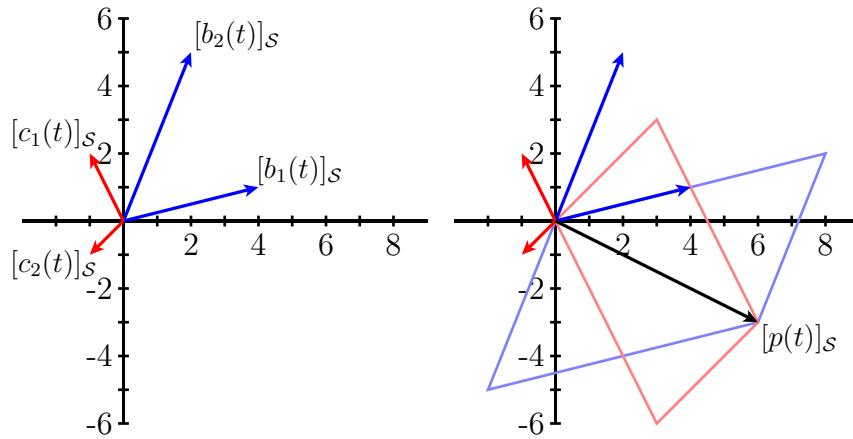


Figure 26.1: Two coordinate systems in \mathbb{R}^2 determined by the bases \mathcal{B} and \mathcal{C} .

Properties of the Change of Basis Matrix

A vector space can have more than one basis, so it is natural to ask how change of bases matrices might be related to one another.

Activity 26.2. The sets $\mathcal{B} = \{3, 4 - t\}$ and $\mathcal{C} = \{1 + 2t, -1 + t\}$ are bases for \mathbb{P}_1 .

- Find the change of basis matrix $P_{\mathcal{C} \leftarrow \mathcal{B}}$ from the basis \mathcal{B} to the basis \mathcal{C} .
- Let $p(t) = 2 + 4t$. Find $[p(t)]_{\mathcal{B}}$ and $[p(t)]_{\mathcal{C}}$.
- Verify by matrix multiplication that $[p(t)]_{\mathcal{C}} = P_{\mathcal{C} \leftarrow \mathcal{B}}[p(t)]_{\mathcal{B}}$.
- Find the change of basis matrix $P_{\mathcal{B} \leftarrow \mathcal{C}}$ from the basis \mathcal{C} to the basis \mathcal{B} .
- Verify by matrix multiplication that $[p(t)]_{\mathcal{B}} = P_{\mathcal{B} \leftarrow \mathcal{C}}[p(t)]_{\mathcal{C}}$.
- How, specifically, are the matrices $P_{\mathcal{C} \leftarrow \mathcal{B}}$ and $P_{\mathcal{B} \leftarrow \mathcal{C}}$ related? (Hint: If you don't see a relationship right away, what is the product of these two matrices?)

Activity 26.2 seems to indicate that the inverse of a change of basis matrix is also a change of basis matrix, which assumes that a change of basis matrix is always invertible. The following theorem provides some properties about change of basis matrices. The proofs are left for the exercises.

Theorem 26.3. Let V be a finite dimensional vector space, and let \mathcal{B} , \mathcal{C} , and \mathcal{S} be bases for V . Then

$$(1) \text{ the change of basis matrix } P_{\mathcal{C} \leftarrow \mathcal{B}} \text{ is invertible,}$$

$$(2) P_{\mathcal{C} \leftarrow \mathcal{B}}^{-1} = P_{\mathcal{B} \leftarrow \mathcal{C}},$$

$$(3) P_{\mathcal{S} \leftarrow \mathcal{C}} P_{\mathcal{C} \leftarrow \mathcal{B}} = P_{\mathcal{S} \leftarrow \mathcal{B}}.$$

Examples

What follows are worked examples that use the concepts from this section.

Example 26.4. Let $\mathcal{B} = \{1 + t^2, 2 - t, 3 + t\}$ and $\mathcal{C} = \{t + 2t^2, 3 + t^2, t\}$. Let $p(t) = 6 - 4t + 8t^2$ and $q(t) = 9 - t + 7t^2$.

- Show that \mathcal{B} and \mathcal{C} are bases for \mathbb{P}_2 .
- Find $[p(t)]_{\mathcal{B}}$ and $[q(t)]_{\mathcal{B}}$.
- Find the change of basis matrix $P_{\mathcal{C} \leftarrow \mathcal{B}}$ from \mathcal{B} to \mathcal{C} .
- Use $P_{\mathcal{C} \leftarrow \mathcal{B}}$ to calculate $[p(t)]_{\mathcal{C}}$ and $[q(t)]_{\mathcal{C}}$. Verify that these are the correct coordinate vectors.

Example Solution.

- Let $\mathcal{S} = \{1, t, t^2\}$ be the standard basis for \mathbb{P}_2 . If $\{[1 + t^2]_{\mathcal{S}}, [2 - t]_{\mathcal{S}}, [3 + t]_{\mathcal{S}}\}$ and $\{[t + 2t^2]_{\mathcal{S}}, [3 + t^2]_{\mathcal{S}}, [t]_{\mathcal{S}}\}$ are bases for \mathbb{R}^3 , then by the coordinate transformation, both \mathcal{B} and \mathcal{C} are bases for \mathbb{P}_2 . Now $[1 + t^2]_{\mathcal{S}} = [1 \ 0 \ 1]^T$, $[2 - t]_{\mathcal{S}} = [2 \ -1 \ 0]^T$, and $[3 + t]_{\mathcal{S}} = [3 \ 1 \ 0]^T$. Technology shows that the reduced row echelon form of $[[1+t]_{\mathcal{S}} \ [2-t]_{\mathcal{S}} \ [3+t]_{\mathcal{S}}]$ is I_3 , so \mathcal{B} is a basis for \mathbb{P}_2 . Similarly, $[t+2t^2]_{\mathcal{S}} = [0 \ 1 \ 2]^T$, $[3+t^2]_{\mathcal{S}} = [0 \ 0 \ 1]^T$, and $[t]_{\mathcal{S}} = [0 \ 1 \ 0]^T$. Technology shows that the reduced row echelon form of $[[t+2t^2]_{\mathcal{S}} \ [3+t^2]_{\mathcal{S}} \ [t]_{\mathcal{S}}]$ is I_3 , so \mathcal{C} is also a basis for \mathbb{P}_2 .
- To find $[p(t)]_{\mathcal{B}}$ and $[q(t)]_{\mathcal{B}}$, we need to write $p(t)$ and $q(t)$ as linear combinations of the basis elements in \mathcal{B} . This is equivalent to writing $[p(t)]_{\mathcal{S}}$ and $[q(t)]_{\mathcal{S}}$ in terms of $[1 + t^2]_{\mathcal{S}}$, $[2 - t]_{\mathcal{S}}$, $[3 + t]_{\mathcal{S}}$. Technology shows that the reduced row echelon form of

$$\left[\begin{array}{ccc|cc} 1 & 2 & 3 & 6 & 9 \\ 0 & -1 & 1 & -4 & -1 \\ 1 & 0 & 0 & 8 & 7 \end{array} \right]$$

is

$$\left[\begin{array}{ccc|cc} 1 & 0 & 0 & 8 & 7 \\ 0 & 1 & 0 & 2 & 1 \\ 0 & 0 & 1 & -2 & 0 \end{array} \right].$$

It follows that

$$[p(t)]_{\mathcal{B}} = [8 \ 2 \ -2]^T \text{ and } [q(t)]_{\mathcal{B}} = [7 \ 1 \ 0]^T.$$

- We know that the change of basis matrix $P_{\mathcal{C} \leftarrow \mathcal{B}}$ from \mathcal{B} to \mathcal{C} can be found by row reducing

$$[[t+2t^2]_{\mathcal{S}} \ [3+t^2]_{\mathcal{S}} \ [t]_{\mathcal{S}} \mid [1+t]_{\mathcal{S}} \ [2-t]_{\mathcal{S}} \ [3+t]_{\mathcal{S}}].$$

Technology shows that the reduced row echelon form of

$$\left[\begin{array}{ccc|ccc} 0 & 3 & 0 & 1 & 2 & 3 \\ 1 & 0 & 1 & 0 & -1 & 1 \\ 2 & 1 & 0 & 1 & 0 & 1 \end{array} \right]$$



is

$$\left[\begin{array}{ccc|ccc} 1 & 0 & 0 & \frac{1}{3} & -\frac{1}{3} & -\frac{1}{2} \\ 0 & 1 & 0 & \frac{1}{3} & \frac{2}{3} & 1 \\ 0 & 0 & 1 & -\frac{1}{3} & -\frac{2}{3} & \frac{3}{2} \end{array} \right].$$

So

$$P_{\mathcal{C} \leftarrow \mathcal{B}} = \frac{1}{6} \begin{bmatrix} 2 & -2 & -3 \\ 2 & 4 & 6 \\ -2 & -4 & 9 \end{bmatrix}.$$

(d) We know that $P_{\mathcal{C} \leftarrow \mathcal{B}}[r(t)]_{\mathcal{B}} = [r(t)]_{\mathcal{C}}$ for any polynomial $r(t)$ in \mathbb{P}_2 . Technology shows that

$$\begin{aligned} [p(t)]_{\mathcal{C}} &= \frac{1}{6} \begin{bmatrix} 2 & -2 & -3 \\ 2 & 4 & 6 \\ -2 & -4 & 9 \end{bmatrix} \begin{bmatrix} 8 \\ 2 \\ -2 \end{bmatrix} = \begin{bmatrix} 3 \\ 2 \\ -7 \end{bmatrix} \\ [q(t)]_{\mathcal{C}} &= \frac{1}{6} \begin{bmatrix} 2 & -2 & -3 \\ 2 & 4 & 6 \\ -2 & -4 & 9 \end{bmatrix} \begin{bmatrix} 7 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 2 \\ 3 \\ -3 \end{bmatrix}. \end{aligned}$$

Since $3(t + 2t^2) + 2(3 + t^2) - 7(t) = p(t)$ and $2(t + 2t^2) + 3(3 + t^2) - 3(t) = q(t)$, we see that we have the correct coordinate vectors.

Example 26.5. Let $D_{2 \times 2}$ be the set of diagonal matrices in $\mathcal{M}_{2 \times 2}$.

(a) Show that the set $\mathcal{B} = \{B_1, B_2\}$, where $B_1 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ and $B_2 = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$ is a basis for $D_{2 \times 2}$.

(b) If $M = \begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix}$ in $D_{2 \times 2}$, find $[M]_{\mathcal{B}}$.

(c) Now show that the set $\mathcal{C} = \{C_1, C_2\}$, where $C_1 = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}$ and $C_2 = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}$ is a basis for $D_{2 \times 2}$. If $M = \begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix}$ in D , find $[M]_{\mathcal{C}}$.

(d) Find the change of basis matrix $P_{\mathcal{C} \leftarrow \mathcal{B}}$. Use the change of basis matrix to verify your result from part (b).

Example Solution.

(a) If $x_1 B_1 + x_2 B_2 = 0$, then equating the 1, 1 entries shows that $x_1 = 0$ and equating the 2, 2 entries gives us $x_2 = 0$. So B_1 and B_2 are linearly independent. If $A = \begin{bmatrix} x_1 & 0 \\ 0 & x_2 \end{bmatrix}$ is in D , then $A = x_1 B_1 + x_2 B_2$ and \mathcal{B} spans $D_{2 \times 2}$. Therefore, \mathcal{B} is a basis for $D_{2 \times 2}$.

(b) Since $M = aB_1 + bB_2$, we have $[M]_{\mathcal{B}} = \begin{bmatrix} a \\ b \end{bmatrix}$.



- (c) Consider the equation $x_1C_1 + x_2C_2 = C$, where C is any matrix in $D_{2 \times 2}$. The coefficient matrix of this system is $A = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}$. Since $\det(A) = -3 \neq 0$, it follows that A is invertible and so the Invertible Matrix Theorem tells us that the only solution to $Ax = \mathbf{0}$ is the trivial solution and that the system $Ax = \mathbf{b}$ is always consistent. Thus, \mathcal{C} is linearly independent and spans $D_{2 \times 2}$. So \mathcal{C} is a basis for $D_{2 \times 2}$.

Let $M = \begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix}$. Then $x_1C_1 + x_2C_2 = M$ if $A \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} a \\ b \end{bmatrix}$. So

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = A^{-1} \begin{bmatrix} a \\ b \end{bmatrix} = -\frac{1}{3} \begin{bmatrix} 1 & -2 \\ -2 & 1 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = -\frac{1}{3} \begin{bmatrix} a - 2b \\ b - 2a \end{bmatrix}.$$

$$\text{So } [M]_{\mathcal{C}} = -\frac{1}{3} \begin{bmatrix} a - 2b \\ b - 2a \end{bmatrix}.$$

- (d) We use the basis \mathcal{B} and row reduce

$$[[C_1]_{\mathcal{B}} \ [C_2]_{\mathcal{B}} \ | \ [B_1]_{\mathcal{B}} \ [B_2]_{\mathcal{B}}] = \left[\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 2 & 2 & 0 & 1 \end{array} \right]$$

to

$$\left[\begin{array}{cc|cc} 1 & 0 & -\frac{1}{3} & \frac{2}{3} \\ 0 & 1 & \frac{2}{3} & -\frac{1}{3} \end{array} \right].$$

So $P_{\mathcal{C} \leftarrow \mathcal{B}} = \begin{bmatrix} -\frac{1}{3} & \frac{2}{3} \\ \frac{2}{3} & -\frac{1}{3} \end{bmatrix}$. Let $M = \begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix}$. Then

$$[M]_{\mathcal{C}} = P_{\mathcal{C} \leftarrow \mathcal{B}} [M]_{\mathcal{B}} = \begin{bmatrix} -\frac{1}{3} & \frac{2}{3} \\ \frac{2}{3} & -\frac{1}{3} \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} -\frac{1}{3}a + \frac{2}{3}b \\ \frac{2}{3}a - \frac{1}{3}b \end{bmatrix}$$

as calculated in part (b).

Summary

- If $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_n\}$ and $\mathcal{C} = \{\mathbf{c}_1, \mathbf{c}_2, \dots, \mathbf{c}_n\}$ are two bases for a vector space V , then the change of basis matrix from \mathcal{B} to \mathcal{C} is the matrix

$$P_{\mathcal{C} \leftarrow \mathcal{B}} = [[\mathbf{b}_1]_{\mathcal{C}} \ [\mathbf{b}_2]_{\mathcal{C}} \ \cdots \ [\mathbf{b}_n]_{\mathcal{C}}]$$

that satisfies

$$[\mathbf{x}]_{\mathcal{C}} = P_{\mathcal{C} \leftarrow \mathcal{B}} [\mathbf{x}]_{\mathcal{B}}$$

for any vector \mathbf{x} in V .

- Change of basis matrices allow us to effectively and efficiently transition from one coordinate system to another.



- To find a change of basis matrix we can use the coordinate transformation to transfer all of our calculations to \mathbb{R}^n . In particular, if $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_n\}$ and $\mathcal{C} = \{\mathbf{c}_1, \mathbf{c}_2, \dots, \mathbf{c}_n\}$ are two bases for a vector space V , then for any basis \mathcal{S} of V the reduced row echelon form of the augmented matrix

$$[[\mathbf{c}_1]_{\mathcal{S}} \ [\mathbf{c}_2]_{\mathcal{S}} \ \cdots \ [\mathbf{c}_n]_{\mathcal{S}} \mid [\mathbf{b}_1]_{\mathcal{S}} \ [\mathbf{b}_2]_{\mathcal{S}} \ \cdots \ [\mathbf{b}_n]_{\mathcal{S}}]$$

is

$$\left[I_n \mid \underset{\mathcal{C} \leftarrow \mathcal{B}}{P} \right].$$

- A change of basis matrix is always invertible. What's more, $\underset{\mathcal{C} \leftarrow \mathcal{B}}{P}^{-1} = \underset{\mathcal{B} \leftarrow \mathcal{C}}{P}$, and $\underset{\mathcal{S} \leftarrow \mathcal{C}}{P} \underset{\mathcal{C} \leftarrow \mathcal{B}}{P} = \underset{\mathcal{S} \leftarrow \mathcal{B}}{P}$ for any bases \mathcal{B} , \mathcal{C} , and \mathcal{S} of a vector space V .

Exercises

- (1) Calculate the change of basis matrix $\underset{\mathcal{C} \leftarrow \mathcal{B}}{P}$ in each of the following cases.

(a) $\mathcal{B} = \{[1 \ 2 \ -1]^T, [-1 \ 1 \ 0]^T, [0 \ 0 \ 1]^T\}$ and $\mathcal{C} = \{[0 \ 1 \ 0]^T, [1 \ -1 \ 1]^T, [0 \ 1 \ 1]^T\}$ in \mathbb{R}^3 .

(b) $\mathcal{B} = \{1 + t^3, t - t^2, t, t^3\}$ and $\mathcal{C} = \{1, t, 1 + t^2, t + t^3\}$ in \mathbb{P}_3 .

(c) $\mathcal{B} = \left\{ \begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix}, \begin{bmatrix} -1 & 0 \\ 1 & 1 \end{bmatrix}, \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} -1 & 1 \\ -1 & 1 \end{bmatrix} \right\}$ and
 $\mathcal{C} = \left\{ \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \right\}$ in $\mathcal{M}_{2 \times 2}$.

- (2) A *permutation matrix* is a change of basis matrix that is obtained when the order of the basis vectors is switched. Let $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3, \mathbf{b}_4\}$ and $\mathcal{C} = \{\mathbf{b}_2, \mathbf{b}_3, \mathbf{b}_1, \mathbf{b}_4\}$ be two ordered bases for a vector space V . Find $\underset{\mathcal{C} \leftarrow \mathcal{B}}{P}$.

(3)

- (a) Suppose $V = \mathbb{P}_1$, $\mathcal{B} = \{1 + 2t, 2 - 3t\}$ and $\mathcal{C} = \{t, 2 - t\}$. You may assume that \mathcal{B} and \mathcal{C} are bases for V . Calculate $\underset{\mathcal{C} \leftarrow \mathcal{B}}{P}$ and $\underset{\mathcal{B} \leftarrow \mathcal{C}}{P}$. How are these matrices related?

- (b) Now let V be an arbitrary n -dimensional vector space and \mathcal{B} and \mathcal{C} arbitrary bases for V . In the remainder of this exercise we demonstrate that the result of part (a) is true in general.

i. Show that $\underset{\mathcal{C} \leftarrow \mathcal{B}}{P}$ is an invertible matrix.

ii. Explain why $\underset{\mathcal{C} \leftarrow \mathcal{B}}{P}^{-1}$ is the change of basis matrix $\underset{\mathcal{B} \leftarrow \mathcal{C}}{P}$.

(4)

- (a) Let $A_1 = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}$, $A_2 = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}$, $A_3 = \begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix}$, $A_4 = \begin{bmatrix} 0 & 2 \\ 1 & 1 \end{bmatrix}$, $A_5 = \begin{bmatrix} 0 & 2 \\ 0 & 1 \end{bmatrix}$, $A_6 = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$. Suppose $V = M_{2 \times 2}$, and that $\mathcal{B} = \{A_1, A_2, A_3, A_4\}$, $\mathcal{C} = \{A_1, A_3, A_4, A_5\}$, and $\mathcal{S} = \{A_2, A_4, A_5, A_6\}$. Calculate each of $\underset{\mathcal{S} \leftarrow \mathcal{C}}{P}$, $\underset{\mathcal{C} \leftarrow \mathcal{B}}{P}$, and $\underset{\mathcal{S} \leftarrow \mathcal{B}}{P}$. How is the product $\underset{\mathcal{S} \leftarrow \mathcal{C}}{P} \underset{\mathcal{C} \leftarrow \mathcal{B}}{P}$ related to $\underset{\mathcal{S} \leftarrow \mathcal{B}}{P}$?

- (b) Now let V be an arbitrary n -dimensional vector space and \mathcal{B} , \mathcal{C} , and \mathcal{S} arbitrary bases for V . Prove that the result of part (a) is true in general. That is, show that

$$\underset{\mathcal{S} \leftarrow \mathcal{C}}{P} \underset{\mathcal{C} \leftarrow \mathcal{B}}{P} = \underset{\mathcal{S} \leftarrow \mathcal{B}}{P}.$$

- (5) We can view the matrix transformation that performs a counterclockwise rotation by an angle θ around the origin in \mathbb{R}^2 as a change of basis matrix. Let $\mathcal{B} = \{\mathbf{e}_1, \mathbf{e}_2\}$ be the standard basis for \mathbb{R}^2 , and let $\mathcal{C} = \{\mathbf{v}_1, \mathbf{v}_2\}$, where $\mathbf{v}_1 = [\cos(\theta) \ \sin(\theta)]^\top$ and $\mathbf{v}_2 = [\cos(\theta + \pi/2) \ \sin(\theta + \pi/2)]^\top$. Note that \mathbf{v}_1 is a vector rotated counterclockwise from the positive x -axis by the angle θ , and \mathbf{v}_2 is a vector rotated counterclockwise from the positive y -axis by the angle θ .

- (a) Use necessary trigonometric identities to show that the change of basis matrix from \mathcal{C} to \mathcal{B} is

$$\begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix}.$$

Then find the change of basis matrix from \mathcal{B} to \mathcal{C} .

- (b) Let $\mathbf{x} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$ in \mathbb{R}^2 . Find $[\mathbf{x}]_{\mathcal{B}}$. Then find $[\mathbf{x}]_{\mathcal{C}}$, where $\mathcal{C} = \{\mathbf{v}_1, \mathbf{v}_2\}$ with $\theta = 30^\circ$.

Draw a picture to illustrate how the components of $[\mathbf{x}]_{\mathcal{C}}$ determine coordinates of $(2, 1)$ in the coordinate system with axes \mathbf{v}_1 and \mathbf{v}_2 .

- (c) Let \mathbf{y} be the vector such that $[\mathbf{y}]_{\mathcal{C}} = [2 \ 3]^\top$. Find $[\mathbf{y}]_{\mathcal{B}}$. Draw a picture to illustrate how the components of $[\mathbf{y}]_{\mathcal{B}}$ determine coordinates of \mathbf{y} in the coordinate system with axes \mathbf{e}_1 and \mathbf{e}_2 .

- (6) Let $\mathcal{B} = \{t, 1+t, t^2\}$ be a basis for \mathbb{P}_2 . Suppose \mathcal{C} is another basis for \mathbb{P}_2 and

$$\underset{\mathcal{C} \leftarrow \mathcal{B}}{P} = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 2 & -1 \\ 2 & 0 & 2 \end{bmatrix}.$$

Find the polynomials in the basis \mathcal{C} .

- (7) Activity 26.1 showed how we can find a change of basis matrix from a basis \mathcal{B} to a basis \mathcal{C} . We can approach this problem in other ways as well. Let \mathcal{B} and \mathcal{C} be bases for a vector space V , and assume that there is some convenient standard basis \mathcal{S} for V .

- (a) Explain why the matrices $\underset{\mathcal{S} \leftarrow \mathcal{B}}{P}$ and $\underset{\mathcal{S} \leftarrow \mathcal{C}}{P}$ will be easy to find.

- (b) How can we combine the matrices $\underset{\mathcal{S} \leftarrow \mathcal{B}}{P}$ and $\underset{\mathcal{S} \leftarrow \mathcal{C}}{P}$ to calculate the change of basis matrix $\underset{\mathcal{C} \leftarrow \mathcal{B}}{P}$? (Hint: There may be inverses involved.)

- (c) Explain why row reducing the augmented matrix $\left[\begin{smallmatrix} P_{\mathcal{S} \leftarrow \mathcal{C}} & | & P_{\mathcal{S} \leftarrow \mathcal{B}} \end{smallmatrix} \right]$ will result in the matrix $\left[\begin{smallmatrix} I_n & | & P_{\mathcal{C} \leftarrow \mathcal{B}} \end{smallmatrix} \right]$.

- (8) Label each of the following statements as True or False. Provide justification for your response. Throughout, the sets $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_n\}$ and $\mathcal{C} = \{\mathbf{c}_1, \mathbf{c}_2, \dots, \mathbf{c}_n\}$ are bases for an n -dimensional vector space V .



- (a) **True/False** The columns of $P_{\mathcal{C} \leftarrow \mathcal{B}}$ span V .
- (b) **True/False** The columns of $P_{\mathcal{C} \leftarrow \mathcal{B}}$ span \mathbb{R}^n .
- (c) **True/False** The columns of $P_{\mathcal{C} \leftarrow \mathcal{B}}$ are linearly independent.
- (d) **True/False** The matrix $[\mathbf{b}_1 \ \mathbf{b}_2 \ \cdots \ \mathbf{b}_n \mid \mathbf{c}_1 \ \mathbf{c}_2 \ \cdots \ \mathbf{c}_n]$ row reduces to $[I_n \mid P_{\mathcal{C} \leftarrow \mathcal{B}}]$.

Project: Planetary Orbits and Change of Basis

We are interested in determining the orbit of planet that orbits the sun. Finding the equation of such an orbit is not difficult, but just having an equation is not enough. For many purposes, it is important to know where the planet is from the perspective or earth observation. This is a more complicated question, one we can address through change of bases matrices.¹

Project Activity 26.1. Since planetary orbits are elliptical, not circular, we need to understand ellipses. An ellipse is a shape like a flattened circle. More specifically, while a circle is the set of points equidistant from a fixed point, an ellipse is a set of points so that the sum of the distances from a point on the ellipse to two fixed points (called foci) is a constant. We can use this definition to derive an equation for an ellipse. We will simplify our work by rotating and translating an ellipse so that its foci are at points $(-c, 0)$ and $(c, 0)$, and the constant sum is $2a$. Let (x, y) be a point on the ellipse as illustrated in Figure 26.2. Use the fact that the sum of the distances from (x, y) to the foci is $2a$ to show that (x, y) satisfies the equation

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1, \quad (26.2)$$

where the points $(0, b)$ and $(0, -b)$ are the y intercepts of the ellipse.

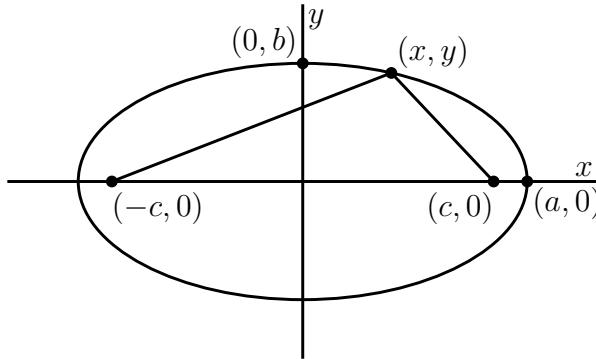


Figure 26.2: An ellipse.

The longer axis of an ellipse is called the major axis and the axis perpendicular to the major axis through the origin is the minor axis. Half of these axes (from the origin) are the semi-major

¹This project is based on the paper “Planetary Orbits: Change of Basis in \mathbb{R}^3 ”, Donald Teets, *Teaching Mathematics and its Applications: An International Journal of the IMA*, Volume 17, Issue 2, 1 June 1998, Pages 66-68.

axis and the semi-minor axis. So the parameter a in (26.2) is the length of the semi-major axis and the parameter b is the length of the semi-minor axis. Note that the points $(0, b)$ and $(0, -b)$ are the y intercepts and the points $(a, 0)$ and $(-a, 0)$ are the x intercepts of this ellipse. Note that if a and b are equal, then the ellipse is a circle. How far the ellipse deviates from a circle is called the *eccentricity* (usually denoted as e) of the ellipse. In other words, the eccentricity is a measure of how flattened an ellipse is, and this is determined by how close c is to a , or how close the ratio $\frac{c}{a}$ is to 1. Thus, we define the eccentricity of an ellipse by

$$e = \frac{c}{a} = \sqrt{1 - \frac{b^2}{a^2}}.$$

Now we assume we have a planet (different from the earth) orbiting the sun and we establish how to convert back and forth from the coordinate system of earth's orbit to the coordinate system of the planet's orbit. To do so we need to establish some coordinate systems. We assume the orbit of earth is in the standard xy plane, with the sun (one of the foci) at the origin. The elliptical orbit of the planet is in some other plane with coordinate axes x' and y' . The two orbital planes intersect in a line. Let this line be the x' axis and let α be the angle the positive x' axis makes with the positive x axis. We can represent the elliptical orbit of the planet in the $x'y'$ plane, but the x' and y' axes are not likely to be the best axes for this orbit. So we define a third coordinate system $x''y''$ in the $x'y'$ plane so that the origin (the position of the sun) is at one focus of the planet's orbit and the x'' axis is the major axis of the orbit and the y'' axis is the minor axis of the orbit of the planet. The unit vectors $\mathbf{b}_1, \mathbf{b}_2$, and \mathbf{b}_3 in the positive x, y , and z directions define a basis $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3\}$ for \mathbb{R}^3 , the unit vectors $\mathbf{b}'_1, \mathbf{b}'_2, \mathbf{b}'_3$ in the positive x', y' , and z' directions define a basis $\mathcal{B}' = \{\mathbf{b}'_1, \mathbf{b}'_2, \mathbf{b}'_3\}$ for \mathbb{R}^3 , and the unit vectors $\mathbf{b}''_1, \mathbf{b}''_2, \mathbf{b}''_3$ in the positive $x'', y'',$ and z'' directions define a basis $\mathcal{B}'' = \{\mathbf{b}''_1, \mathbf{b}''_2, \mathbf{b}''_3\}$ for \mathbb{R}^3 . See Figure 26.3 for illustrations. Finally, let γ be the angle between the

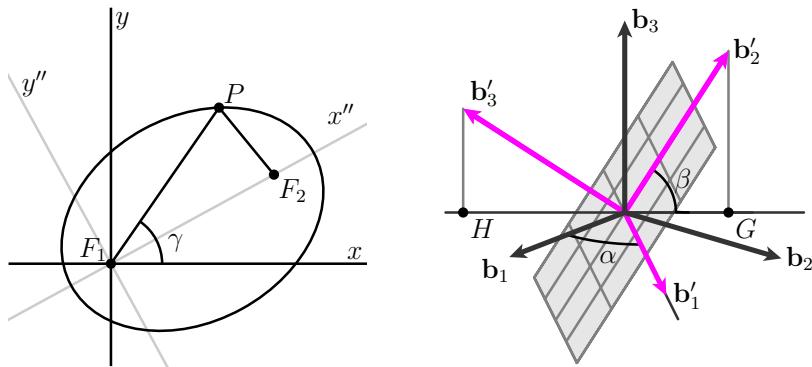


Figure 26.3: Left: The planet's orbit in the $x''y''$ system. Right: The planes of the planet and earth orbits.

positive x' axis and the positive x'' axis as shown at left in Figure 26.3. Our first step is to find the change of basis matrix from \mathcal{B}'' to \mathcal{B}' .

Project Activity 26.2. Explain why the change of basis matrix $P_{\mathcal{B}' \leftarrow \mathcal{B}''}$ is given by

$$P_{\mathcal{B}' \leftarrow \mathcal{B}''} = \begin{bmatrix} \cos(\gamma) & -\sin(\gamma) & 0 \\ \sin(\gamma) & \cos(\gamma) & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

More complicated is the change of basis matrix from \mathcal{B}' to \mathcal{B} .

Project Activity 26.3. Now we look for $P_{\mathcal{B} \leftarrow \mathcal{B}'} = [[\mathbf{b}'_1]_{\mathcal{B}} \ [\mathbf{b}'_2]_{\mathcal{B}} \ [\mathbf{b}'_3]_{\mathcal{B}}]$. Assume that the plane p in which the planet's orbit lies has equation $z = ax + by$.

- (a) Explain why $\mathbf{b}'_1 = \cos(\alpha)\mathbf{b}_1 + \sin(\alpha)\mathbf{b}_2$.
- (b) The x' axis is the intersection of the plane $z = ax + by$ with the plane $z = 0$, so the equation of x' axis in terms of x and y is $ax + by = 0$. Now we determine the coordinates of \mathbf{b}'_2 in terms of the basis \mathcal{B} .
 - i. Explain why the vector $[b - a 0]^T$ lies on the x' axis. We take this vector to point in the positive x' direction. This gives us another representation of \mathbf{b}'_1 – namely that $\mathbf{b}'_1 = \frac{1}{\sqrt{a^2+b^2}}[b - a 0]^T$.
 - ii. Explain why a vector in the plane $z = ax + by$ orthogonal to \mathbf{b}'_1 is $[a b a^2 + b^2]^T$.
 - iii. From the previous part we have

$$\mathbf{b}'_2 = \frac{1}{\sqrt{a^2 + b^2 + (a^2 + b^2)^2}} [b \ a \ a^2 + b^2]^T.$$

Let $G = \left(\frac{b}{\sqrt{a^2+b^2+(a^2+b^2)^2}}, \frac{a}{\sqrt{a^2+b^2+(a^2+b^2)^2}}, 0 \right)$ be the terminal point of the projection of \mathbf{b}'_2 onto the xy plane. Show that \overrightarrow{OG} is orthogonal to \mathbf{b}'_1 .

- iv. Let β be the angle between the plane p and the xy plane as illustrated at right in Figure 26.3. Explain why $\|\overrightarrow{OG}\| = \cos(\beta)$. Then explain why

$$\mathbf{b}'_2 = [-\cos(\beta) \sin(\alpha) \ \cos(\beta) \cos(\alpha) \ \sin(\beta)]^T.$$

(Hint: Use the trigonometric identities $\cos(A + \frac{\pi}{2}) = -\sin(A)$ and $\sin(A + \frac{\pi}{2}) = \cos(A)$.)

- (c) Finally, we find $[\mathbf{b}'_3]_{\mathcal{B}}$. The cross product of \mathbf{b}'_1 and \mathbf{b}'_2 is a vector orthogonal to \mathbf{b}'_1 and \mathbf{b}'_2 , so

$$\mathbf{b}'_3 = \frac{1}{\sqrt{(a^2 + b^2)^2 + 4a^2b^2}} [-a(a^2 + b^2) \ -b(a^2 + b^2) \ 2ab]^T.$$

Let H be the terminal point of the projection of \mathbf{b}'_3 onto the xy plane as illustrated at right in Figure 26.3.

- i. Explain why the angle between \mathbf{b}'_1 and \overrightarrow{OH} is $\frac{\pi}{2}$.
- ii. Explain why $\|\overrightarrow{OH}\| = \sin(\beta)$. (Hint: Use the trigonometric identity $\cos(\frac{\pi}{2} - A) = \sin(A)$.)
- iii. Since the angle from \mathbf{b}_1 to \overrightarrow{OH} is negative, this angle is $\alpha - \frac{\pi}{2}$. Use this angle and the previous information to find the coordinates of the point H and, consequently, explain why

$$\mathbf{b}'_3 = [\sin(\beta) \sin(\alpha) \ -\sin(\beta) \cos(\alpha) \ \cos(\beta)]^T.$$

(Hint: Use the trigonometric identities $\cos(A - \frac{\pi}{2}) = \sin(A)$ and $\sin(A - \frac{\pi}{2}) = -\cos(A)$.)



- (d) Explain why the change of basis matrix $P_{\mathcal{B} \leftarrow \mathcal{B}'}^P$ from \mathcal{B}' to \mathcal{B} is

$$P_{\mathcal{B} \leftarrow \mathcal{B}'}^P = \begin{bmatrix} \cos(\alpha) & -\cos(\beta)\sin(\alpha) & \sin(\beta)\sin(\alpha) \\ \sin(\alpha) & \cos(\beta)\cos(\alpha) & -\sin(\beta)\cos(\alpha) \\ 0 & \sin(\beta) & \cos(\beta) \end{bmatrix}.$$

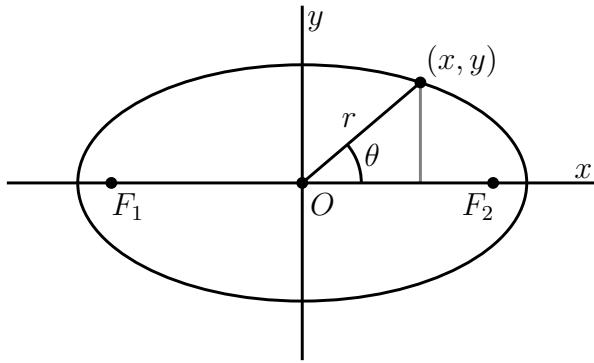


Figure 26.4: Points on the ellipse in terms of angles.

With the change of basis matrices we can convert from any one coordinate system to the other. Note that all of the change of basis matrices are written in terms of angles, so it will be convenient to have a way to express points on our ellipses using angles as well. Given any point on an ellipse (or any point in the plane), we can represent the coordinates of that point in terms of the angle θ the vector through the origin and the point makes with the positive x -axis and the distance r from the origin to the point as shown in Figure 26.4. In this representation we have $x = r \cos(\theta)$ and $y = r \sin(\theta)$.

So we can start in the $x''y''$ coordinate system with the coordinate vector of a point $[\overrightarrow{OP}]_{\mathcal{B}''} = [r \cos(\theta) \ r \sin(\theta) \ 0]^T$. Then to view this point in the xy system, we apply the change of basis matrices

$$[\overrightarrow{OP}]_{\mathcal{B}} = P_{\mathcal{B} \leftarrow \mathcal{B}'}^P P_{\mathcal{B}' \leftarrow \mathcal{B}''}^P [r \cos(\theta) \ r \sin(\theta) \ 0]^T.$$

Of course we can also convert from \mathcal{B} coordinates to \mathcal{B}'' coordinates by applying the inverses of our change of basis matrices.



Part VI

Inner Product Spaces

Section 27

The Dot Product in \mathbb{R}^n

Focus Questions

By the end of this section, you should be able to give precise and thorough answers to the questions listed below. You may want to keep these questions in mind to focus your thoughts as you complete the section.

- What is the dot product of two vectors? Under what conditions is the dot product defined?
- How do we find the angle between two nonzero vectors in \mathbb{R}^n ?
- How does the dot product tell us if two vectors are orthogonal?
- How do we define the length of a vector in any dimension and how can the dot product be used to calculate the length?
- How do we define the distance between two vectors?
- What is the orthogonal projection of a vector \mathbf{u} in the direction of the vector \mathbf{v} and how do we find it?
- What is the orthogonal complement of a subspace W of \mathbb{R}^n ?

Application: Hidden Figures in Computer Graphics

In video games, the speed at which a computer can render changing graphics views is vitally important. To increase a computer's ability to render a scene, programs often try to identify those parts of the images a viewer could see and those parts the viewer could not see. For example, in a scene involving buildings, the viewer could not see any images blocked by a solid building. In the mathematical world, this can be visualized by graphing surfaces. In Figure 27.1 we see a crude image of a house made up of small polygons (this is how programs generally represent surfaces). On the left in Figure 27.1 we see all of the polygons that are needed to construct the entire surface, even those polygons that lie behind others which we could not see if the surface was solid. On the

right in Figure 27.1 we have hidden the parts of the polygons that we cannot see from our view. We

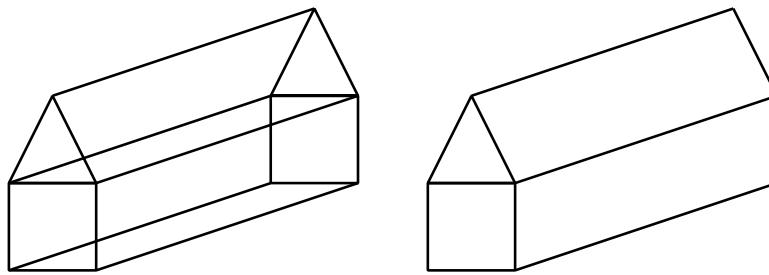


Figure 27.1: Images of a house.

also see this idea in mathematics when we graph surfaces. Figure 27.2 shows the graph of the surface defined by $f(x, y) = \sqrt{4 - x^2}$ that is made up of polygons. At left we see all of the polygons and at right only those parts that would be visible from our viewing perspective. By eliminating

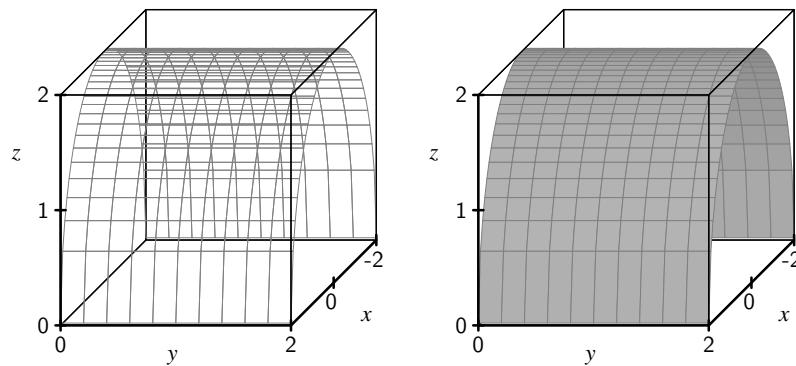


Figure 27.2: Graphs of $f(x, y) = \sqrt{4 - x^2}$.

the parts of the polygons we cannot see from our viewing perspective, the computer program can more quickly render the viewing image. Later in this section we will explore one method for how programs remove the hidden portions of images. This process involves the dot product of vectors.

Introduction

Orthogonality, a concept which generalizes the idea of perpendicularity, is an important concept in linear algebra. We use the dot product to define orthogonality and more generally angles between vectors in \mathbb{R}^n for any dimension n . The dot product has many applications, e.g., finding components of forces acting in different directions in physics and engineering. The dot product is also an example of a larger concept, *inner products*, that we will discuss a bit later. We introduce and investigate dot products in this section.

We will illustrate the dot product in \mathbb{R}^2 , but the process we go through will translate to any

dimension. Recall that we can represent the vector $\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$ as the directed line segment (or arrow) from the origin to the point (v_1, v_2) in \mathbb{R}^2 , as illustrated in Figure 27.3. Using the Pythagorean Theorem we can then define the length (or magnitude or norm) of the vector \mathbf{v} in \mathbb{R}^2 as

$$\|\mathbf{v}\| = \sqrt{v_1^2 + v_2^2}.$$

We can also write this norm as

$$\sqrt{v_1 v_1 + v_2 v_2}.$$

The expression under the square root is an important one and we extend it and give it a special name.

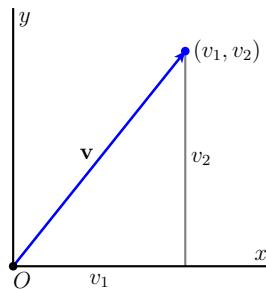


Figure 27.3: A vector in \mathbb{R}^2 from the origin to a point.

If $\mathbf{u} = [u_1 \ u_2]^\top$ and $\mathbf{v} = [v_1 \ v_2]^\top$ are vectors in \mathbb{R}^2 , then we call the expression $u_1 v_1 + u_2 v_2$ the *dot product* of \mathbf{u} and \mathbf{v} , and denote it as $\mathbf{u} \cdot \mathbf{v}$. With this idea in mind, we can rewrite the norm of the vector \mathbf{v} as

$$\|\mathbf{v}\| = \sqrt{\mathbf{v} \cdot \mathbf{v}}.$$

The definition of the dot product translates naturally to \mathbb{R}^n (see Exercise 5 in Section 5).

Definition 27.1. Let $\mathbf{u} = [u_1 \ u_2 \ \cdots \ u_n]$ and $\mathbf{v} = [v_1 \ v_2 \ \cdots \ v_n]$ be vectors in \mathbb{R}^n . The **dot product** (or **scalar product**) of \mathbf{u} and \mathbf{v} is the scalar

$$\mathbf{u} \cdot \mathbf{v} = u_1 v_1 + u_2 v_2 + \cdots + u_n v_n = \sum_{i=1}^n u_i v_i.$$

The dot product then allows us to define the norm (or magnitude or length) of any vector in \mathbb{R}^n .

Definition 27.2. The **norm** $\|\mathbf{v}\|$ of the vector $\mathbf{v} \in \mathbb{R}^n$ is

$$\|\mathbf{v}\| = \sqrt{\mathbf{v} \cdot \mathbf{v}}.$$

We also use the words **magnitude** or **length** as alternatives for the word norm. We can equivalently write the norm of the vector $\mathbf{v} = [v_1 \ v_2 \ \cdots \ v_n]^\top$ as

$$\|\mathbf{v}\| = \sqrt{v_1^2 + v_2^2 + \cdots + v_n^2}.$$

We can also realize the dot product as a matrix product. If $\mathbf{u} = [u_1 \ u_2 \ \cdots \ u_n]^T$ and $\mathbf{v} = [v_1 \ v_2 \ \cdots \ v_n]^T$, then

$$\mathbf{u} \cdot \mathbf{v} = \mathbf{u}^T \mathbf{v}.^1$$

IMPORTANT NOTE: The dot product is only defined between two vectors with the *same number of components*.

Preview Activity 27.1.

- (1) Find $\mathbf{u} \cdot \mathbf{v}$ if $\mathbf{u} = [2 \ 3 \ -1 \ 4]^T$ and $\mathbf{v} = [4 \ 6 \ 7 \ -5]^T$ in \mathbb{R}^4 .
- (2) The dot product satisfies some useful properties as given in the next theorem.

Theorem 27.3. *Let \mathbf{u} , \mathbf{v} , and \mathbf{w} be vectors in \mathbb{R}^n , and let c be a scalar. Then*

- (a) $\mathbf{u} \cdot \mathbf{v} = \mathbf{v} \cdot \mathbf{u}$ (the dot product is commutative),
- (b) $(\mathbf{u} + \mathbf{v}) \cdot \mathbf{w} = (\mathbf{u} \cdot \mathbf{w}) + (\mathbf{v} \cdot \mathbf{w})$ (the dot product distributes over vector addition),
- (c) $(c\mathbf{u}) \cdot \mathbf{v} = \mathbf{u} \cdot (c\mathbf{v}) = c(\mathbf{u} \cdot \mathbf{v})$,
- (d) $\mathbf{u} \cdot \mathbf{u} \geq 0$ with equality if and only if $\mathbf{u} = \mathbf{0}$,
- (e) $\|c\mathbf{u}\| = |c|\|\mathbf{u}\|$.

Verification of some of these properties is left to the exercises. Let \mathbf{u} and \mathbf{v} be vectors in \mathbb{R}^5 with $\mathbf{u} \cdot \mathbf{v} = -1$, $\|\mathbf{u}\| = 2$ and $\|\mathbf{v}\| = 3$. Use the properties of the dot product given in Theorem 27.3 to find each of the following.

- (a) $\mathbf{u} \cdot 2\mathbf{v}$
- (b) $(\mathbf{u} + \mathbf{v}) \cdot \mathbf{v}$
- (c) $(2\mathbf{u} + 4\mathbf{v}) \cdot (\mathbf{u} - 7\mathbf{v})$
- (3) At times we will want to find vectors in the direction of a given vector that have a certain magnitude. Let $\mathbf{u} = [2 \ 2 \ 1]^T$ in \mathbb{R}^3 .
 - (a) What is $\|\mathbf{u}\|$?
 - (b) Show that $\left\| \frac{1}{\|\mathbf{u}\|} \mathbf{u} \right\| = 1$.
 - (c) Vectors with magnitude 1 are important and are given a special name.

Definition 27.4. A vector \mathbf{v} in \mathbb{R}^n is a **unit vector** if $\|\mathbf{v}\| = 1$.

We can use unit vectors to find vectors of a given length in the direction of a given vector. Let c be a positive scalar and \mathbf{v} a vector in \mathbb{R}^n . Use properties from Theorem 27.3 to show that the magnitude of the vector $c \frac{\mathbf{v}}{\|\mathbf{v}\|}$ is c .

¹Technically, $\mathbf{u}^T \mathbf{v}$ is a 1×1 matrix and not a scalar, but we usually think of 1×1 matrices as scalars.



The Distance Between Vectors

Finding optimal solutions to systems is an important problem in applied mathematics. It is often the case that we cannot find an exact solution that satisfies certain constraints, so we look instead for the “best” solution that satisfies the constraints. An example of this is fitting a least squares line to a set of data. As we will see, the dot product (and inner products in general) will allow us to find “best” solutions to certain types of problems, where we measure accuracy using the notion of a distance between vectors. Geometrically, we can represent a vector \mathbf{u} as a directed line segment from the origin to the point defined by \mathbf{u} . If we have two vectors \mathbf{u} and \mathbf{v} , we can think of the length of the difference $\mathbf{u} - \mathbf{v}$ as a measure of how far apart the two vectors are from each other. It is natural, then to define the distance between vectors as follows.

Definition 27.5. Let \mathbf{u} and \mathbf{v} be vectors in \mathbb{R}^n . The **distance** between \mathbf{u} and \mathbf{v} is the length of the difference $\mathbf{u} - \mathbf{v}$ or

$$\|\mathbf{u} - \mathbf{v}\|.$$

As Figure 27.4 illustrates, if vectors \mathbf{u} and \mathbf{v} emanate from the same initial point, and P and Q are the terminal points of \mathbf{u} and \mathbf{v} , respectively, then the difference $\|\mathbf{u} - \mathbf{v}\|$ is the standard Euclidean distance between the points P and Q .

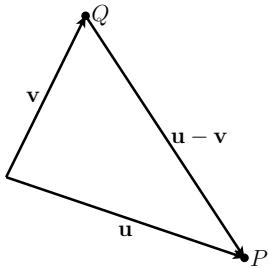
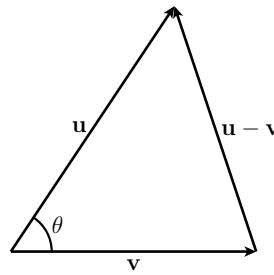


Figure 27.4: $\|\mathbf{u} - \mathbf{v}\|$.

The Angle Between Two Vectors

Determining a “best” solution to a problem often involves finding a solution that minimizes a distance. We generally accomplish a minimization through orthogonality – which depends on the angle between vectors. Given two vectors \mathbf{u} and \mathbf{v} in \mathbb{R}^n , we position the vectors so that they emanate from the same initial point. If the vectors are nonzero, then they determine a plane in \mathbb{R}^n . In that plane there are two angles that these vectors create. We will define the angle between the vectors to be the smaller of these two angles. The dot product will tell us how to find the angle between vectors. Let \mathbf{u} and \mathbf{v} be vectors in \mathbb{R}^n and θ the angle between them as illustrated in Figure 27.5. Using the Law of Cosines, we have

$$\|\mathbf{u} - \mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 - 2\|\mathbf{u}\| \|\mathbf{v}\| \cos(\theta).$$

Figure 27.5: The angle between \mathbf{u} and \mathbf{v}

By rearranging, we obtain

$$\begin{aligned} \|\mathbf{u}\| \|\mathbf{v}\| \cos(\theta) &= \frac{1}{2} (\|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 - \|\mathbf{u} - \mathbf{v}\|^2) \\ &= \frac{1}{2} (\|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 - (\mathbf{u} - \mathbf{v}) \cdot (\mathbf{u} - \mathbf{v})) \\ &= \frac{1}{2} (\|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 - \mathbf{u} \cdot \mathbf{u} + 2\mathbf{u} \cdot \mathbf{v} - \mathbf{v} \cdot \mathbf{v}) \\ &= \mathbf{u} \cdot \mathbf{v}. \end{aligned}$$

So the angle θ between two nonzero vectors \mathbf{u} and \mathbf{v} in \mathbb{R}^n satisfies the equation

$$\cos(\theta) = \frac{\mathbf{u} \cdot \mathbf{v}}{\|\mathbf{u}\| \|\mathbf{v}\|}. \quad (27.1)$$

Of particular interest to us will be the situation where vectors \mathbf{u} and \mathbf{v} are *orthogonal* (perpendicular).² Intuitively, we think of two vectors as orthogonal if the angle between them is 90° .

Activity 27.1.

- (a) The vectors $\mathbf{e}_1 = [1 \ 0]^\top$ and $\mathbf{e}_2 = [0 \ 1]^\top$ are perpendicular in \mathbb{R}^2 . What is $\mathbf{e}_1 \cdot \mathbf{e}_2$?
 - (b) Now let \mathbf{u} and \mathbf{v} be any vectors in \mathbb{R}^n .
 - i. Suppose the angle between nonzero vectors \mathbf{u} and \mathbf{v} is 90° . What does Equation (27.1) tell us about $\mathbf{u} \cdot \mathbf{v}$?
 - ii. Now suppose that $\mathbf{u} \cdot \mathbf{v} = 0$. What does Equation (27.1) tell us about the angle between \mathbf{u} and \mathbf{v} ? Why?
 - iii. Explain why the following definition makes sense.
- Definition 27.6.** Two vectors \mathbf{u} and \mathbf{v} in \mathbb{R}^n are **orthogonal** if $\mathbf{u} \cdot \mathbf{v} = 0$.
- iv. According to Definition 27.6, to which vectors is $\mathbf{0}$ orthogonal? Does this make sense to you intuitively? Explain.

Activity 27.2.

²We use the term orthogonal instead of perpendicular because we will be able to extend this idea to situations where we normally don't think of objects as being perpendicular.

- (a) Find the angle between the two vectors $\mathbf{u} = [1 \ 3 \ -2 \ 5]^\top$ and $\mathbf{v} = [5 \ 2 \ 3 \ -1]^\top$.

(b) Find, if possible, two non-parallel vectors orthogonal to $\mathbf{u} = \begin{bmatrix} 0 \\ 3 \\ -2 \\ 1 \end{bmatrix}$.

Orthogonal Projections

When running a sprint, the racers may be aided or slowed by the wind. The wind assistance is a measure of the wind speed that is helping push the runners down the track. It is much easier to run a very fast race if the wind is blowing hard in the direction of the race. So that world records aren't dependent on the weather conditions, times are only recorded as record times if the wind aiding the runners is less than or equal to 2 meters per second. Wind speed for a race is recorded by a wind gauge that is set up close to the track. It is important to note, however, that weather is not always as cooperative as we might like. The wind does not always blow exactly in the direction of the track, so the gauge must account for the angle the wind makes with the track. If the wind is blowing in the direction of the vector \mathbf{u} in Figure 27.6 and the track is in the direction of the vector \mathbf{v} in Figure 27.6, then only part of the total wind vector is actually working to help the runners. This part is called the orthogonal projection of the vector \mathbf{u} onto the vector \mathbf{v} and is denoted $\text{proj}_{\mathbf{v}} \mathbf{u}$. The next activity shows how to find this projection.

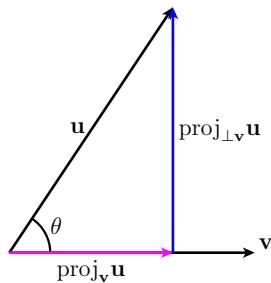


Figure 27.6: The orthogonal projection of \mathbf{u} onto \mathbf{v} .

Activity 27.3. Since the orthogonal projection $\text{proj}_{\mathbf{v}} \mathbf{u}$ is in the direction of \mathbf{v} , there exists a constant c such that $\text{proj}_{\mathbf{v}} \mathbf{u} = c\mathbf{v}$. If we solve for c , we can find $\text{proj}_{\mathbf{v}} \mathbf{u}$.

- The wind component that acts perpendicular to the direction of \mathbf{v} is called the projection of \mathbf{u} orthogonal to \mathbf{v} and is denoted $\text{proj}_{\perp \mathbf{v}} \mathbf{u}$ as shown in Figure 27.6. Write an equation that involves $\text{proj}_{\mathbf{v}} \mathbf{u}$, $\text{proj}_{\perp \mathbf{v}} \mathbf{u}$, and \mathbf{u} . Then solve that equation for $\text{proj}_{\perp \mathbf{v}} \mathbf{u}$.
- Given that \mathbf{v} and $\text{proj}_{\perp \mathbf{v}} \mathbf{u}$ are orthogonal, what does that tell us about $\mathbf{v} \cdot \text{proj}_{\perp \mathbf{v}} \mathbf{u}$? Combine this fact with the result of part (a) and that $\text{proj}_{\mathbf{v}} \mathbf{u} = c\mathbf{v}$ to obtain an equation involving \mathbf{v} , \mathbf{u} , and c .
- Solve for c using the equation you found in the previous step.
- Use your value of c to identify $\text{proj}_{\mathbf{v}} \mathbf{u}$.

To summarize:

Definition 27.7. Let \mathbf{u} and \mathbf{v} be vectors in \mathbb{R}^n with $\mathbf{v} \neq \mathbf{0}$.

- (1) The **orthogonal projection** of \mathbf{u} onto \mathbf{v} is the vector

$$\text{proj}_{\mathbf{v}} \mathbf{u} = \frac{\mathbf{u} \cdot \mathbf{v}}{\|\mathbf{v}\|^2} \mathbf{v}. \quad (27.2)$$

- (2) The **projection** of \mathbf{u} orthogonal to \mathbf{v} is the vector

$$\text{proj}_{\perp \mathbf{v}} \mathbf{u} = \mathbf{u} - \text{proj}_{\mathbf{v}} \mathbf{u}.$$

Activity 27.4. Let $\mathbf{u} = \begin{bmatrix} 5 \\ 8 \end{bmatrix}$ and $\mathbf{v} = \begin{bmatrix} 6 \\ -10 \end{bmatrix}$. Find $\text{proj}_{\mathbf{v}} \mathbf{u}$ and $\text{proj}_{\perp \mathbf{v}} \mathbf{u}$ and draw a picture to illustrate.

Orthogonal Complements

In Activity 27.1 we defined two vectors \mathbf{u} and \mathbf{v} in \mathbb{R}^n to be orthogonal (or perpendicular) if $\mathbf{u} \cdot \mathbf{v} = 0$. With this in mind we can define the orthogonal complement of a subspace of \mathbb{R}^n .

Definition 27.8. Let W be a subspace of \mathbb{R}^n for some $n \geq 1$. The **orthogonal complement** of W is the set

$$W^\perp = \{\mathbf{x} \in \mathbb{R}^n : \mathbf{x} \cdot \mathbf{w} = 0 \text{ for all } \mathbf{w} \in W\}.$$

Preview Activity 27.2. Let $W = \text{Span}\{[1 \ -1]^\top\}$ in \mathbb{R}^2 . Completely describe all vectors in W^\perp both algebraically and geometrically.

The orthogonal projection of a vector \mathbf{u} onto a vector \mathbf{v} is really a projection of the vector \mathbf{u} onto the vector space $\text{Span}\{\mathbf{v}\}$. The vector $\text{proj}_{\mathbf{v}} \mathbf{u}$ is the best approximation to \mathbf{u} of all the vectors in $\text{Span}\{\mathbf{v}\}$ in the sense that $\text{proj}_{\mathbf{v}} \mathbf{u}$ is the closest to \mathbf{u} among all vectors in $\text{Span}\{\mathbf{v}\}$, as we will prove later. Another familiar example where we see this type of behavior is when we look at planes in 3-space. Remember that a plane through the origin in \mathbb{R}^3 is a two dimensional subspace of \mathbb{R}^3 . We define a plane through the origin to be the set of all vectors in \mathbb{R}^3 that are orthogonal to a given vector (called a *normal* vector). For example, to find the equation of the plane through the origin in \mathbb{R}^3 orthogonal to the normal vector $\mathbf{n} = [1 \ 2 \ -1]^\top$, we seek all the vectors $\mathbf{v} = [x \ y \ z]^\top$ such that

$$\mathbf{v} \cdot \mathbf{n} = 0.$$

This gives us the equation

$$x + 2y - z = 0$$

as the equation of this plane.

There is a more general idea here as defined in Preview Activity 27.2. If we have a set S of vectors in \mathbb{R}^n , we let S^\perp (read as “ S perp”, called the *orthogonal complement* of S) be the set of all vectors in \mathbb{R}^n that are orthogonal to every vector in S . In our plane example, the set S is $\{\mathbf{n}\}$ and S^\perp is the plane with equation $x + 2y - z = 0$.



Activity 27.5. We have seen another example of orthogonal complements. Let A be an $m \times n$ matrix with rows $\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_m$ in order. Consider the three spaces $\text{Nul } A$, $\text{Row } A$, and $\text{Col } A$ related to A , where $\text{Row } A = \text{Span}\{\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_m\}$ (that is, $\text{Row } A$ is the span of the rows of A). Let \mathbf{x} be a vector in $\text{Row } A$.

- What does it mean for \mathbf{x} to be in $\text{Row } A$?
- Now let \mathbf{y} be a vector in $\text{Nul } A$. Use the result of part (a) and the fact that $A\mathbf{y} = \mathbf{0}$ to explain why $\mathbf{x} \cdot \mathbf{y} = 0$. (Hint: Calculate $A\mathbf{y}$ using scalar products of rows of A with \mathbf{y} .) Explain how this verifies $(\text{Row } A)^\perp = \text{Nul } A$.
- Use A^\top in place of A in the result of the previous part to show $(\text{Col } A)^\perp = \text{Nul } A^\top$.

The activity proves the following theorem:

Theorem 27.9. *Let A be an $m \times n$ matrix. Then*

$$(\text{Row } A)^\perp = \text{Nul } A \text{ and } (\text{Col } A)^\perp = \text{Nul } A^\top.$$

To show that a vector is in the orthogonal complement of a subspace, it is not necessary to demonstrate that the vector is orthogonal to every vector in the subspace. If we have a basis for the subspace, it suffices to show that the vector is orthogonal to every vector in that basis for the subspace, as the next theorem demonstrates.

Theorem 27.10. *Let $\mathcal{B} = \{\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_m\}$ be a basis for a subspace W of \mathbb{R}^n . A vector \mathbf{v} in \mathbb{R}^n is orthogonal to every vector in W if and only if \mathbf{v} is orthogonal to every vector in \mathcal{B} .*

Proof. Let $\mathcal{B} = \{\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_m\}$ be a basis for a subspace W of \mathbb{R}^n and let \mathbf{v} be a vector in \mathbb{R}^n . Our theorem is a biconditional, so we need to prove both implications. Since $\mathcal{B} \subset W$, it follows that if \mathbf{v} is orthogonal to every vector in W , then \mathbf{v} is orthogonal to every vector in \mathcal{B} . This proves the forward implication. Now we assume that \mathbf{v} is orthogonal to every vector in \mathcal{B} and show that \mathbf{v} is orthogonal to every vector in W . Let \mathbf{x} be a vector in W . Then

$$\mathbf{x} = x_1\mathbf{w}_1 + x_2\mathbf{w}_2 + \cdots + x_m\mathbf{w}_m$$

for some scalars x_1, x_2, \dots, x_m . Then

$$\begin{aligned} \mathbf{v} \cdot \mathbf{x} &= \mathbf{v} \cdot (x_1\mathbf{w}_1 + x_2\mathbf{w}_2 + \cdots + x_m\mathbf{w}_m) \\ &= x_1(\mathbf{v} \cdot \mathbf{w}_1) + x_2(\mathbf{v} \cdot \mathbf{w}_2) + \cdots + x_m(\mathbf{v} \cdot \mathbf{w}_m) \\ &= 0. \end{aligned}$$

Thus, \mathbf{v} is orthogonal to \mathbf{x} and \mathbf{v} is orthogonal to every vector in W . ■

Activity 27.6. Let $W = \text{Span} \left\{ \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\}$. Find all vectors in W^\perp .

We will work more closely with projections and orthogonal complements in later sections.



Examples

What follows are worked examples that use the concepts from this section.

Example 27.11. Let ℓ be the line defined by the equation $ax + by + c = 0$ with in \mathbb{R}^2 and let $P = (x_0, y_0)$ be a point in the plane. In this example we will learn how to find the distance from P to ℓ .

- (a) Show that $\mathbf{n} = [a \ b]^\top$ is orthogonal to the line ℓ . That is, \mathbf{n} is orthogonal to any vector on the line ℓ .
- (b) Let $Q = (x_1, y_1)$ be any point on line ℓ . Draw a representative picture of P , \mathbf{n} with its initial point at P , along with Q and ℓ . Explain how to use a projection to determine the distance from P to ℓ .
- (c) Use the idea from part (b) to show that the distance d from P to ℓ satisfies

$$d = \frac{|ax_0 + by_0 + c|}{\sqrt{a^2 + b^2}}. \quad (27.3)$$

- (d) Use Equation (27.3) to find the distance from the point $(3, 4)$ to the line $y = 2x + 1$.

Example Solution.

- (a) Any vector on the line ℓ is a vector between two points on the line. Let $Q = (x_1, y_1)$ and $R = (x_2, y_2)$ be points on the line ℓ . Then $\mathbf{u} = \overrightarrow{QR} = [x_2 - x_1 \ y_2 - y_1]^\top$ is a vector on line ℓ . Since Q and R are on the line, we know that $ax_1 + by_1 + c = 0$ and $ax_2 + by_2 + c = 0$. So $-c = ax_1 + by_1 = ax_2 + by_2$ and

$$0 = a(x_2 - x_1) + b(y_2 - y_1) = [a \ b]^\top \mathbf{u}.$$

Thus, $\mathbf{n} = [a \ b]^\top$ is orthogonal to every vector on the line ℓ .

- (b) A picture of the situation is shown in Figure 27.7. If $\mathbf{v} = \overrightarrow{PQ}$, then the distance from point P to line ℓ is given by $\|\text{proj}_{\mathbf{n}} \mathbf{v}\|$.

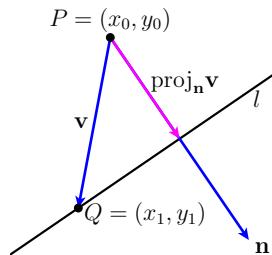


Figure 27.7: Distance from a point to a line.

- (c) Recall that $\mathbf{n} = [a \ b]^T$ and $\mathbf{v} = \overrightarrow{PQ} = [x_1 - x_0 \ y_1 - y_0]^T$. Since $ax_1 + by_1 + c = 0$, we have

$$\begin{aligned}\text{proj}_{\mathbf{n}} \mathbf{v} &= \frac{\mathbf{v} \cdot \mathbf{n}}{\|\mathbf{n}\|^2} \mathbf{n} \\ &= \frac{a(x_1 - x_0) + b(y_1 - y_0)}{a^2 + b^2} [a \ b]^T \\ &= \frac{ax_1 + by_1 - ax_0 - by_0}{a^2 + b^2} [a \ b]^T \\ &= \frac{-c - ax_0 - by_0}{a^2 + b^2} [a \ b]^T.\end{aligned}$$

So

$$\|\text{proj}_{\mathbf{n}} \mathbf{v}\| = \frac{|ax_0 + by_0 + c|}{\sqrt{a^2 + b^2}} \sqrt{a^2 + b^2} = \frac{|ax_0 + by_0 + c|}{\sqrt{a^2 + b^2}}.$$

- (d) Here we have $P = (3, 4)$, and the equation of our line is $2x - y + 1 = 0$. So $a = 2$, $b = -1$, and $c = -1$. Thus, the distance from P to the line is

$$\frac{|ax_0 + by_0 + c|}{\sqrt{a^2 + b^2}} = \frac{|2(3) - (4) + 1|}{\sqrt{4 + 1}} = \frac{3}{\sqrt{5}}.$$

Example 27.12. Let a , b , and c be scalars with $a \neq 0$, and let

$$W = \{ax + by + cz = 0 : x, y, z \in \mathbb{R}\}.$$

- (a) Find two vectors that span W , showing that W is a subspace of \mathbb{R}^3 . (In fact, W is a plane through the origin in \mathbb{R}^3 .)
- (b) Find a vector \mathbf{n} that is orthogonal to the two vectors you found in part (a).
- (c) Explain why $\{\mathbf{n}\}$ is a basis for W^\perp .

Example Solution.

- (a) The coefficient matrix for the system $ax + by + cz = 0$ is $[a \ b \ c]^T$. The first column is a pivot column and the others are not. So y and z are free variables and

$$[x \ y \ z]^T = \left[-\frac{b}{a}y - \frac{c}{a}z, y, z \right]^T = y \left[-\frac{b}{a} \ 1 \ 0 \right]^T + z \left[-\frac{c}{a} \ 0 \ 1 \right]^T.$$

$$\text{So } W = \text{Span} \left\{ \left[-\frac{b}{a} \ 1 \ 0 \right]^T, \left[-\frac{c}{a} \ 0 \ 1 \right]^T \right\}.$$

- (b) If we let $\mathbf{n} = [a \ b \ c]^T$, then

$$\begin{aligned}\mathbf{n} \cdot \left[-\frac{b}{a} \ 1 \ 0 \right]^T &= -b + b = 0 \\ \mathbf{n} \cdot \left[-\frac{c}{a} \ 0 \ 1 \right]^T &= -c + c = 0.\end{aligned}$$

Thus, $[a \ b \ c]^T$ is orthogonal to both $\left[-\frac{b}{a} \ 1 \ 0 \right]^T$ and $\left[-\frac{c}{a} \ 0 \ 1 \right]^T$.



- (c) Let $\mathbf{u} = \left[-\frac{b}{a} \ 1 \ 0 \right]^\top$ and $\mathbf{v} = \left[-\frac{c}{a} \ 0 \ 1 \right]^\top$. Every vector in W has the form $x\mathbf{u} + y\mathbf{v}$ for some scalars x and y , and

$$\mathbf{n} \cdot (x\mathbf{u} + y\mathbf{v}) = x(\mathbf{n} \cdot \mathbf{u}) + y(\mathbf{n} \cdot \mathbf{v}) = 0.$$

So $\mathbf{n} \in W^\perp$.

Now we need to verify that $\{\mathbf{n}\}$ spans W^\perp . Let $\mathbf{w} = [w_1 \ w_2 \ w_3]^\top$ be in W^\perp . Then $\mathbf{w} \cdot \mathbf{z} = 0$ for every $\mathbf{z} \in W$. In particular, $\mathbf{w} \cdot \mathbf{u} = 0$ or $-\frac{b}{a}w_1 + w_2 = 0$, and $\mathbf{w} \cdot \mathbf{v} = 0$ or $-\frac{c}{a}w_1 + w_3 = 0$. Equivalently, we have $w_2 = \frac{b}{a}w_1$ and $w_3 = \frac{c}{a}w_1$. So

$$\begin{aligned}\mathbf{w} &= [w_1 \ w_2 \ w_3]^\top \\ &= \left[w_1 \ \frac{b}{a}w_1 \ \frac{c}{a}w_1 \right]^\top \\ &= \frac{1}{a}[a \ b \ c]^\top w_1 \\ &= \frac{w_1}{a}\mathbf{n}.\end{aligned}$$

So every vector in W^\perp is a multiple of \mathbf{n} , and $\{\mathbf{n}\}$ spans W^\perp . We conclude that $\{\mathbf{n}\}$ is a basis for W^\perp . Thus, the vector $[a \ b \ c]^\top$ is a normal vector to the plane $ax + by + cz = 0$ if $a \neq 0$. The same reasoning works if at least one of a , b , or c is nonzero, so we can say in every case that $[a \ b \ c]^\top$ is a normal vector to the plane $ax + by + cz = 0$.

Summary

- The dot product of vectors $\mathbf{u} = [u_1 \ u_2 \ \cdots \ u_n]^\top$ and $\mathbf{v} = [v_1 \ v_2 \ \cdots \ v_n]^\top$ in \mathbb{R}^n is the scalar

$$\mathbf{u} \cdot \mathbf{v} = u_1v_1 + u_2v_2 + \cdots + u_nv_n = \sum_{i=1}^n u_i v_i.$$

- The angle θ between two nonzero vectors \mathbf{u} and \mathbf{v} in \mathbb{R}^n satisfies the equation

$$\cos(\theta) = \frac{\mathbf{u} \cdot \mathbf{v}}{\|\mathbf{u}\| \|\mathbf{v}\|}$$

and $0 \leq \theta \leq 180$.

- Two vectors \mathbf{u} and \mathbf{v} are orthogonal if $\mathbf{u} \cdot \mathbf{v} = 0$.
- The length, or norm, of the vector \mathbf{u} can be found as $\|\mathbf{u}\| = \sqrt{\mathbf{u} \cdot \mathbf{u}}$.
- The distance between the vectors \mathbf{u} and \mathbf{v} in \mathbb{R}^n is $\|\mathbf{u} - \mathbf{v}\|$, which is the length of the difference $\mathbf{u} - \mathbf{v}$.
- Let \mathbf{u} and \mathbf{v} be vectors in \mathbb{R}^n .
 - The orthogonal projection of \mathbf{u} onto \mathbf{v} is the vector

$$\text{proj}_{\mathbf{v}} \mathbf{u} = \frac{\mathbf{u} \cdot \mathbf{v}}{\|\mathbf{v}\|^2} \mathbf{v}.$$



- The projection of \mathbf{u} perpendicular to \mathbf{v} is the vector

$$\text{proj}_{\perp \mathbf{v}} \mathbf{u} = \mathbf{u} - \text{proj}_{\mathbf{v}} \mathbf{u}.$$

- The orthogonal complement of the subspace W of \mathbb{R}^n is the set

$$W^\perp = \{\mathbf{x} \in \mathbb{R}^n : \mathbf{x} \cdot \mathbf{w} = 0 \text{ for all } \mathbf{w} \in W\}.$$

Exercises

- (1) For each of the following pairs of vectors, find $\mathbf{u} \cdot \mathbf{v}$, calculate the angle between \mathbf{u} and \mathbf{v} , determine if \mathbf{u} and \mathbf{v} are orthogonal, find $\|\mathbf{u}\|$ and $\|\mathbf{v}\|$, calculate the distance between \mathbf{u} and \mathbf{v} , and determine the orthogonal projection of \mathbf{u} onto \mathbf{v} .
 - (a) $\mathbf{u} = [1 \ 2]^\top, \mathbf{v} = [-2 \ 1]^\top$
 - (b) $\mathbf{u} = [2 \ -2]^\top, \mathbf{v} = [1 \ -1]^\top$
 - (c) $\mathbf{u} = [2 \ -1]^\top, \mathbf{v} = [1 \ 3]^\top$
 - (d) $\mathbf{u} = [1 \ 2 \ 0]^\top, \mathbf{v} = [-2 \ 1 \ 1]^\top$
 - (e) $\mathbf{u} = [0 \ 0 \ 1]^\top, \mathbf{v} = [1 \ 1 \ 1]^\top$
- (2) Given $\mathbf{u} = [2 \ 1 \ 2]^\top$, find a vector \mathbf{v} so that the angle between \mathbf{u} and \mathbf{v} is 60° and the orthogonal projection of \mathbf{v} onto \mathbf{u} has length 2.
- (3) For which value(s) of h is the angle between $[1 \ 1 \ h]^\top$ and $[1 \ 2 \ 1]^\top$ equal to 60° ?
- (4) Let $A = [a_{ij}]$ be a $k \times m$ matrix with rows $\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_k$, and let $B = [\mathbf{b}_1 \ \mathbf{b}_2 \ \cdots \ \mathbf{b}_n]$ be an $m \times n$ matrix with columns $\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_n$. Show that we can write the matrix product AB in a shorthand way as $AB = [\mathbf{r}_i \cdot \mathbf{b}_j]$.
- (5) Let A be an $m \times n$, \mathbf{u} a vector in \mathbb{R}^n and \mathbf{v} a vector in \mathbb{R}^m . Show that

$$A\mathbf{u} \cdot \mathbf{v} = \mathbf{u} \cdot A^\top \mathbf{v}.$$

- (6) Let \mathbf{u}, \mathbf{v} , and \mathbf{w} be vectors in \mathbb{R}^n . Show that
 - (a) $(\mathbf{u} + \mathbf{v}) \cdot \mathbf{w} = (\mathbf{u} \cdot \mathbf{w}) + (\mathbf{v} \cdot \mathbf{w})$ (the dot product *distributes over vector addition*)
 - (b) If c is an arbitrary constant, then $(c\mathbf{u}) \cdot \mathbf{v} = \mathbf{u} \cdot (c\mathbf{v}) = c(\mathbf{u} \cdot \mathbf{v})$
- (7) The Pythagorean Theorem states that if a and b are the lengths of the legs of a right triangle whose hypotenuse has length c , then $a^2 + b^2 = c^2$. If we think of the legs as defining vectors \mathbf{u} and \mathbf{v} , then the hypotenuse is the vector $\mathbf{u} + \mathbf{v}$ and we can restate the Pythagorean Theorem as

$$\|\mathbf{u} + \mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2.$$

In this exercise we show that this result holds in any dimension.

- (a) Let \mathbf{u} and \mathbf{v} be orthogonal vectors in \mathbb{R}^n . Show that $\|\mathbf{u} + \mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2$.
(Hint: Rewrite $\|\mathbf{u} + \mathbf{v}\|^2$ using the dot product.)



- (b) Must it be true that if \mathbf{u} and \mathbf{v} are vectors in \mathbb{R}^n with $\|\mathbf{u} + \mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2$, then \mathbf{u} and \mathbf{v} are orthogonal? If not, provide a counterexample. If true, verify the statement.

- (8) The Cauchy-Schwarz inequality,

$$|\mathbf{u} \cdot \mathbf{v}| \leq \|\mathbf{u}\| \|\mathbf{v}\| \quad (27.4)$$

for any vectors \mathbf{u} and \mathbf{v} in \mathbb{R}^n , is considered one of the most important inequalities in mathematics. We verify the Cauchy-Schwarz inequality in this exercise. Let \mathbf{u} and \mathbf{v} be vectors in \mathbb{R}^n .

- (a) Explain why the inequality (27.4) is true if either \mathbf{u} or \mathbf{v} is the zero vector. As a consequence, we assume that \mathbf{u} and \mathbf{v} are nonzero vectors for the remainder of this exercise.

- (b) Let $\mathbf{w} = \text{proj}_{\mathbf{v}} \mathbf{u} = \frac{\mathbf{u} \cdot \mathbf{v}}{\|\mathbf{v}\|^2} \mathbf{v}$ and let $\mathbf{z} = \mathbf{u} - \mathbf{w}$. We know that $\mathbf{w} \cdot \mathbf{z} = 0$. Use Exercise 7. of this section to show that

$$\|\mathbf{u}\|^2 \geq \|\mathbf{w}\|^2.$$

- (c) Now show that $\|\mathbf{w}\|^2 = \frac{|\mathbf{u} \cdot \mathbf{v}|^2}{\|\mathbf{v}\|^2}$.

- (d) Combine parts (b) and (c) to explain why equation (27.4) is true.

- (9) Let \mathbf{u} and \mathbf{v} be vectors in \mathbb{R}^n . Then \mathbf{u} , \mathbf{v} and $\mathbf{u} + \mathbf{v}$ form a triangle. We should then expect that the length of any one side of the triangle is smaller than the sum of the lengths of the other sides (since the straight line distance is the shortest distance between two points). In other words, we expect that

$$\|\mathbf{u} + \mathbf{v}\| \leq \|\mathbf{u}\| + \|\mathbf{v}\|. \quad (27.5)$$

Equation (27.5) is called the *Triangle Inequality*. Use the Cauchy-Schwarz inequality (Exercise 8) to prove the triangle inequality.

- (10) Let W be a subspace of \mathbb{R}^n for some n . Show that W^\perp is also a subspace of \mathbb{R}^n .
- (11) Let W be a subspace of \mathbb{R}^n . Show that W is a subspace of $(W^\perp)^\perp$.
- (12) If W is a subspace of \mathbb{R}^n for some n , what is $W \cap W^\perp$? Verify your answer.
- (13) Suppose $W_1 \subseteq W_2$ are two subspaces of \mathbb{R}^n . Show that $W_2^\perp \subseteq W_1^\perp$.
- (14) Label each of the following statements as True or False. Provide justification for your response.
- (a) **True/False** The dot product is defined between any two vectors.
 - (b) **True/False** If \mathbf{u} and \mathbf{v} are vectors in \mathbb{R}^n , then $\mathbf{u} \cdot \mathbf{v}$ is another vector in \mathbb{R}^n .
 - (c) **True/False** If \mathbf{u} and \mathbf{v} are vectors in \mathbb{R}^n , then $\mathbf{u} \cdot \mathbf{v}$ is always non-negative.
 - (d) **True/False** If \mathbf{v} is a vector in \mathbb{R}^n , then $\mathbf{v} \cdot \mathbf{v}$ is never negative.
 - (e) **True/False** If \mathbf{u} and \mathbf{v} are vectors in \mathbb{R}^n and $\mathbf{u} \cdot \mathbf{v} = 0$, then $\mathbf{u} = \mathbf{v} = \mathbf{0}$.



- (f) **True/False** If \mathbf{v} is a vector in \mathbb{R}^n and $\mathbf{v} \cdot \mathbf{v} = 0$, then $\mathbf{v} = \mathbf{0}$.
- (g) **True/False** The norm of the sum of vectors is the sum of the norms of the vectors.
- (h) **True/False** If \mathbf{u} and \mathbf{v} are vectors in \mathbb{R}^n , then $\text{proj}_{\mathbf{v}} \mathbf{u}$ is a vector in the same direction as \mathbf{u} .
- (i) **True/False** The only subspace W of \mathbb{R}^n for which $W^\perp = \{\mathbf{0}\}$ is $W = \mathbb{R}^n$.
- (j) **True/False** If a vector \mathbf{u} is orthogonal to \mathbf{v}_1 and \mathbf{v}_2 , then \mathbf{u} is also orthogonal to $\mathbf{v}_1 + \mathbf{v}_2$.
- (k) **True/False** If a vector \mathbf{u} is orthogonal to \mathbf{v}_1 and \mathbf{v}_2 , then \mathbf{u} is also orthogonal to all linear combinations of \mathbf{v}_1 and \mathbf{v}_2 .
- (l) **True/False** If $\mathbf{u} \neq \mathbf{0}$ and \mathbf{v} are parallel, then the orthogonal projection of \mathbf{v} onto \mathbf{u} equals \mathbf{v} .
- (m) **True/False** If $\mathbf{u} \neq \mathbf{0}$ and \mathbf{v} are orthogonal, then the orthogonal projection of \mathbf{v} onto \mathbf{u} equals \mathbf{v} .
- (n) **True/False** For any vector \mathbf{v} and $\mathbf{u} \neq \mathbf{0}$, $\|\text{proj}_{\mathbf{u}} \mathbf{v}\| \leq \|\mathbf{v}\|$.
- (o) **True/False** Given an $m \times n$ matrix, $\dim(\text{Row } A) + \dim(\text{Row } A)^\perp = n$.
- (p) **True/False** If A is a square matrix, then the columns of A are orthogonal to the vectors in $\text{Nul } A$.
- (q) **True/False** The vectors in the null space of an $m \times n$ matrix are orthogonal to vectors in the row space of A .

Project: Back-Face Culling

To identify hidden polygons in a surface, we will utilize a technique called *back face culling*. This involves identifying which polygons are back facing and which are front facing relative to the viewer's perspective. The first step is to assign a direction to each polygon in a surface.

Project Activity 27.1. Consider the polygon $ABCD$ in Figure 27.8. Since a polygon is flat, every vector in the polygon is perpendicular to a fixed vector (which we call a *normal vector* to the polygon). A normal vector \mathbf{n} for the polygon $ABCD$ in Figure 27.8 is shown. In this activity we learn how to find a normal vector to a polygon.

Let $\mathbf{x} = [x_1 \ x_2 \ x_3]^\top$ and $\mathbf{y} = [y_1 \ y_2 \ y_3]^\top$ be two vectors in \mathbb{R}^3 . If \mathbf{x} and \mathbf{y} are linearly independent, then \mathbf{x} and \mathbf{y} determine a polygon as shown in Figure 27.8. Our goal is to find a vector \mathbf{n} that is orthogonal to both \mathbf{x} and \mathbf{y} . Let $\mathbf{w} = [w_1 \ w_2 \ w_3]^\top$ be another vector in \mathbb{R}^3 and let $C = \begin{bmatrix} \mathbf{w}^\top \\ \mathbf{x}^\top \\ \mathbf{y}^\top \end{bmatrix}$ be the matrix whose rows are \mathbf{w} , \mathbf{x} , and \mathbf{y} . Let C_{ij} be the ij th cofactor of C , that is C_{ij} is $(-1)^{i+j}$ times the determinant of the submatrix of C obtained by deleting the i th row and j th column of C . Now define the vector $\mathbf{x} \times \mathbf{y}$ as follows:

$$\mathbf{x} \times \mathbf{y} = C_{11}\mathbf{e}_1 + C_{12}\mathbf{e}_2 + C_{13}\mathbf{e}_3.$$



The vector $\mathbf{x} \times \mathbf{y}$ is called the *cross product* of the vectors \mathbf{x} and \mathbf{y} . (Note that the cross product is only defined for vectors in \mathbb{R}^3 .) We will show that $\mathbf{x} \times \mathbf{y}$ is orthogonal to both \mathbf{x} and \mathbf{y} , making $\mathbf{x} \times \mathbf{y}$ a normal vector to the polygon defined by \mathbf{x} and \mathbf{y} .

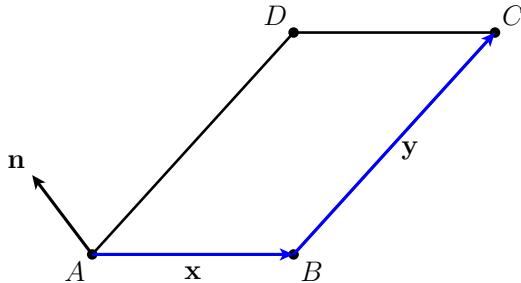


Figure 27.8: Normal vector to a polygon.

- (a) Show that

$$\mathbf{x} \times \mathbf{y} = \begin{bmatrix} x_2y_3 - x_3y_2 \\ x_3y_1 - x_1y_3 \\ x_1y_2 - x_2y_1 \end{bmatrix}.$$

- (b) Use a cofactor expansion of C along the first row and properties of the dot product to show that

$$\det(C) = \mathbf{w} \cdot (\mathbf{x} \times \mathbf{y}).$$

- (c) Use the result of part (b) and properties of the determinant to calculate $\mathbf{x} \cdot (\mathbf{x} \times \mathbf{y})$ and $\mathbf{y} \cdot (\mathbf{x} \times \mathbf{y})$. Explain why $\mathbf{x} \times \mathbf{y}$ is orthogonal to both \mathbf{x} and \mathbf{y} and is therefore a normal vector to the polygon determined by \mathbf{x} and \mathbf{y} .

Project Activity 27.1 shows how we can find a normal vector to a parallelogram – take two vectors \mathbf{x} and \mathbf{y} between the vertices of the parallelogram and calculate their cross products. Such a normal vector can define a direction for the parallelogram. There is still a problem, however.

Project Activity 27.2. Let $\mathbf{x} = [x_1 \ x_2 \ x_3]^\top$ and $\mathbf{y} = [y_1 \ y_2 \ y_3]^\top$ be any vectors in \mathbb{R}^3 . There is a relationship between $\mathbf{x} \times \mathbf{y}$ and $\mathbf{y} \times \mathbf{x}$. Find and verify this relationship.

Project Activity 27.2 shows that the cross product is anticommutative, so we get different directions if we switch the order in which we calculate the cross product. To fix a direction, we establish the convention that we always label the vertices of our parallelogram in the counterclockwise direction as shown in Figure 27.8. This way we always use \mathbf{x} as the vector from vertex A to vertex B rather than the reverse. With this convention established, we can now define the direction of a parallelogram as the direction of its normal vector.

Once we have a normal vector established for each polygon, we can now determine which polygons are back-face and which are front-face. Figure 27.9 at left provides the gist of the idea, where we represent the polygons with line segments to illustrate. If the viewer's eye is at point P and views the figures, the normal vectors of the visible polygons point in a direction toward the viewer (front-face) and the normal vectors of the polygons hidden from the viewer point away from the viewer (back-face). What remains is to determine an effective computational way to identify the front and back facing polygons.

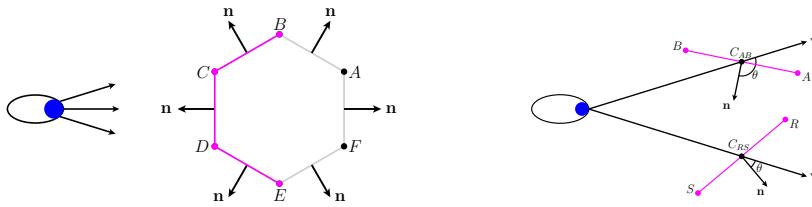


Figure 27.9: Left: Hidden faces. Right: Back face culling.

Project Activity 27.3. Consider the situation as depicted at right in Figure 27.9. Assume that AB and RS are polygons (rendered one dimensionally here) with normal vectors \mathbf{n} at their centers as shown. The viewer's eye is at point P and the viewer's line of vision to the centers C_{AB} and C_{RS} are indicated by the vectors \mathbf{v} . Each vector \mathbf{v} makes an angle θ with the normal to the polygon.

- What can be said about the angle θ for a front-facing polygon? What must be true about $\mathbf{v} \cdot \mathbf{n}$ for a front-facing polygon? Why?
- What can be said about the angle θ for a back-facing polygon? What must be true about $\mathbf{v} \cdot \mathbf{n}$ for a back-facing polygon? Why?
- The dot product then provides us with a simple computational tool for identifying back-facing polygons (assuming we have already calculated all of the normal vectors). We can then create an algorithm to cull the back-facing polygons. Assuming that we the viewpoint P and the coordinates of the polygons of the surface, complete the pseudo-code for a back-face culling algorithm:

```

for all polygons on the surface do
    calculate the normal vector  $\mathbf{n}$  using the _____ product for the current polygon
    calculate the center  $C$  of the current polygon
    calculate the viewing vector _____
    if _____ then
        render the current polygon
    end if
end for

```

As a final comment, back-face culling generally reduces the number of polygons to be rendered by half. This algorithm is not perfect and does not always do what we want it to do (e.g., it may not remove all parts of a polygon that we don't see), so there are other algorithms to use in concert with back-face culling to correctly render objects.

Section 28

Orthogonal and Orthonormal Bases in \mathbb{R}^n

Focus Questions

By the end of this section, you should be able to give precise and thorough answers to the questions listed below. You may want to keep these questions in mind to focus your thoughts as you complete the section.

- What is an orthogonal set in \mathbb{R}^n ? What is one useful fact about orthogonal subsets of \mathbb{R}^n ?
- What is an orthogonal basis for a subspace of \mathbb{R}^n ? What is an orthonormal basis?
- How does orthogonality help us find the weights to write a vector as a linear combination of vectors in an orthogonal basis?
- What is an orthogonal matrix and why are orthogonal matrices useful?

Application: Rotations in 3D

An aircraft in flight, like a plane or the space shuttle, can perform three independent rotations: *roll*, *pitch*, and *yaw*. Roll is a rotation about the axis through the nose and tail of the aircraft, pitch is rotation moving the nose of the aircraft up or down through the axis from wingtip to wingtip, and yaw is the rotation when the nose of the aircraft turns left or right about the axis though the plane from top to bottom. These rotations take place in 3-space and the axes of the rotations change as the aircraft travels through space. To understand how aircraft maneuver, it is important to know about general rotations in space. These are more complicated than rotations in 2-space, and, as we will see later in this section, involve orthogonal sets.

Introduction

If $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_m\}$ is a basis for a subspace W of \mathbb{R}^n , we know that any vector \mathbf{w} in W can be written uniquely as a linear combination of the vectors in \mathcal{B} . In the past, the way we have found the coordinates of \mathbf{x} with respect to \mathcal{B} , i.e. the weights needed to write a vector \mathbf{x} as a linear combination of the elements in \mathcal{B} , has been to row reduce the matrix $[\mathbf{v}_1 \ \mathbf{v}_2 \ \cdots \ \mathbf{v}_n \ | \ \mathbf{x}]$ to solve the corresponding system. This can be a cumbersome process, especially if we need to do it many times. For certain types of bases, namely the *orthogonal* and *orthonormal* bases, there is a much easier way to find the individual weights for this linear combination.

Recall that two nonzero vectors \mathbf{u} and \mathbf{v} in \mathbb{R}^n are orthogonal if $\mathbf{u} \cdot \mathbf{v} = 0$. We can extend this idea to an entire set. For example, the standard basis $\mathcal{S} = \{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ for \mathbb{R}^3 has the property that any two distinct vectors in \mathcal{S} are orthogonal to each other. The basis vectors in \mathcal{S} make a very nice coordinate system for \mathbb{R}^3 , where the basis vectors provide the directions for the coordinate axes. We could rotate this standard basis, or multiply any of the vectors in the basis by a nonzero constant, and retain a basis in which all distinct vectors are orthogonal to each other (e.g., $\{[2 \ 0 \ 0]^\top, [0 \ 3 \ 0]^\top, [0 \ 0 \ 1]^\top\}$). We define this idea of having all vectors be orthogonal to each other for sets, and then later for bases.

Definition 28.1. A non-empty subset S of \mathbb{R}^n is **orthogonal** if $\mathbf{u} \cdot \mathbf{v} = 0$ for every pair of distinct vector \mathbf{u} and \mathbf{v} in S .

Preview Activity 28.1.

- (1) Determine if the set $S = \{[1 \ 2 \ 1]^\top, [2 \ -1 \ 0]^\top\}$ is an orthogonal set.
- (2) Orthogonal bases are especially important.

Definition 28.2. An **orthogonal basis** \mathcal{B} for a subspace W of \mathbb{R}^n is a basis of W that is also an orthogonal set.

Let $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$, where $\mathbf{v}_1 = [1 \ 2 \ 1]^\top$, $\mathbf{v}_2 = [2 \ -1 \ 0]^\top$, and $\mathbf{v}_3 = [1 \ 2 \ -5]^\top$.

- (a) Explain why \mathcal{B} is an orthogonal basis for \mathbb{R}^3 .
- (b) Suppose \mathbf{x} has coordinates x_1, x_2, x_3 with respect to the basis \mathcal{B} , i.e.

$$\mathbf{x} = x_1 \mathbf{v}_1 + x_2 \mathbf{v}_2 + x_3 \mathbf{v}_3.$$

Substitute for \mathbf{x} in $\mathbf{x} \cdot \mathbf{v}_1$ and use the orthogonality property of the basis \mathcal{B} to show that $x_1 = \frac{\mathbf{x} \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1}$. Then determine x_2 and x_3 similarly. Finally, calculate the values of x_1 , x_2 , and x_3 if $\mathbf{x} = [1 \ 1 \ 1]^\top$.

- (c) Find components of $\mathbf{x} = [1 \ 1 \ 1]^\top$ by reducing the augmented matrix $[\mathbf{v}_1 \ \mathbf{v}_2 \ \mathbf{v}_3 \ | \ \mathbf{x}]$. Does this result agree with your work from the previous part?

Orthogonal Sets

We defined orthogonal sets in \mathbb{R}^n and bases of subspaces of \mathbb{R}^n in Definitions 28.1 and 28.2. We saw that the standard basis in \mathbb{R}^3 is an orthogonal set and an orthogonal basis of \mathbb{R}^3 , and there are many other examples as well.



Activity 28.1. Let $\mathbf{w}_1 = \begin{bmatrix} -2 \\ 1 \\ -1 \end{bmatrix}$, $\mathbf{w}_2 = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$, and $\mathbf{w}_3 = \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}$. In the same manner as in Preview Activity 28.1, we can show that the set $S_1 = \{\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3\}$ is an orthogonal subset of \mathbb{R}^3 .

- (a) Is the set $S_2 = \left\{ \begin{bmatrix} -2 \\ 1 \\ -1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix} \right\}$ an orthogonal subset of \mathbb{R}^3 ?
- (b) Suppose a vector \mathbf{v} is a vector so that $S_1 \cup \{\mathbf{v}\}$ is an orthogonal subset of \mathbb{R}^3 . Then $\mathbf{w}_i \cdot \mathbf{v} = 0$ for each i . Explain why this implies that \mathbf{v} is in $\text{Nul } A$, where $A = \begin{bmatrix} -2 & 1 & -1 \\ 0 & 1 & 1 \\ 1 & 1 & -1 \end{bmatrix}$.
- (c) Assuming that the reduced row echelon form of the matrix A is I_3 , explain why it is not possible to find a nonzero vector \mathbf{v} so that $S_1 \cup \{\mathbf{v}\}$ is an orthogonal subset of \mathbb{R}^3 .

The example from Activity 28.1 suggests that we can have three orthogonal nonzero vectors in \mathbb{R}^3 , but no more. Orthogonal vectors are, in a sense, as far apart as they can be. So we might expect that there is no linear relationship between orthogonal vectors. The following theorem makes this clear.

Theorem 28.3. *Let $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_m\}$ be a set of nonzero orthogonal vectors in \mathbb{R}^n . Then the vectors $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_m$ are linearly independent.*

Proof. Let $S = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_m\}$ be a set of nonzero orthogonal vectors in \mathbb{R}^n . To show that $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_m$ are linearly independent, assume that

$$x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + \cdots + x_m\mathbf{v}_m = \mathbf{0} \quad (28.1)$$

for some scalars x_1, x_2, \dots, x_m . We will show that $x_i = 0$ for each i from 1 to m . Since the vectors in S are orthogonal to each other, we know that $\mathbf{v}_i \cdot \mathbf{v}_j = 0$ whenever $i \neq j$. Fix an index k between 1 and m . We evaluate the dot product of both sides of (28.1) with \mathbf{v}_k and simplify using the dot product properties:

$$\begin{aligned} \mathbf{v}_k \cdot (x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + \cdots + x_m\mathbf{v}_m) &= \mathbf{v}_k \cdot \mathbf{0} \\ (\mathbf{v}_k \cdot x_1\mathbf{v}_1) + (\mathbf{v}_k \cdot x_2\mathbf{v}_2) + \cdots + (\mathbf{v}_k \cdot x_m\mathbf{v}_m) &= 0 \\ x_1(\mathbf{v}_k \cdot \mathbf{v}_1) + x_2(\mathbf{v}_k \cdot \mathbf{v}_2) + \cdots + x_m(\mathbf{v}_k \cdot \mathbf{v}_m) &= 0. \end{aligned} \quad (28.2)$$

Now all of the dot products on the left are 0 except for $\mathbf{v}_k \cdot \mathbf{v}_k$, so (28.2) becomes

$$x_k(\mathbf{v}_k \cdot \mathbf{v}_k) = 0.$$

We assumed that $\mathbf{v}_k \neq \mathbf{0}$ and since $\mathbf{v}_k \cdot \mathbf{v}_k = \|\mathbf{v}_k\|^2 \neq 0$, we conclude that $x_k = 0$. We chose k arbitrarily, so we have shown that $x_k = 0$ for each k between 1 and m . Therefore, the only solution to equation (28.1) is the trivial solution with $x_1 = x_2 = \cdots = x_m = 0$ and the set S is linearly independent. ■



Properties of Orthogonal Bases

Orthogonality is a useful and important property for a basis to have. In Preview Activity 28.1 we saw that if a vector \mathbf{x} in the span of an orthogonal basis $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ could be written as a linear combination of the basis vectors as $\mathbf{x} = x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + x_3\mathbf{v}_3$, then $x_1 = \frac{\mathbf{x} \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1}$. If we continued that same argument we could show that

$$\mathbf{x} = \left(\frac{\mathbf{x} \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1} \right) \mathbf{v}_1 + \left(\frac{\mathbf{x} \cdot \mathbf{v}_2}{\mathbf{v}_2 \cdot \mathbf{v}_2} \right) \mathbf{v}_2 + \left(\frac{\mathbf{x} \cdot \mathbf{v}_3}{\mathbf{v}_3 \cdot \mathbf{v}_3} \right) \mathbf{v}_3.$$

We can apply this idea in general to see how the orthogonality of an orthogonal basis allows us to quickly and easily determine the weights to write a given vector as a linear combination of orthogonal basis vectors. To see why, let $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_m\}$ be an orthogonal basis for a subspace W of \mathbb{R}^n and let \mathbf{x} be any vector in W . We know that

$$\mathbf{x} = x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + \cdots + x_m\mathbf{v}_m$$

for some scalars x_1, x_2, \dots, x_m . Let $1 \leq k \leq m$. Then, using orthogonality of vectors $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_m$, we have

$$\mathbf{v}_k \cdot \mathbf{x} = x_1(\mathbf{v}_k \cdot \mathbf{v}_1) + x_2(\mathbf{v}_k \cdot \mathbf{v}_2) + \cdots + x_m(\mathbf{v}_k \cdot \mathbf{v}_m) = x_k \mathbf{v}_k \cdot \mathbf{v}_k.$$

So

$$x_k = \frac{\mathbf{x} \cdot \mathbf{v}_k}{\mathbf{v}_k \cdot \mathbf{v}_k}.$$

Thus, we can calculate each weight individually with two simple dot products. We summarize this discussion in the next theorem.

Theorem 28.4. *Let $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_m\}$ be an orthogonal basis for a subspace of \mathbb{R}^n . Let \mathbf{x} be a vector in W . Then*

$$\mathbf{x} = \frac{\mathbf{x} \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1} \mathbf{v}_1 + \frac{\mathbf{x} \cdot \mathbf{v}_2}{\mathbf{v}_2 \cdot \mathbf{v}_2} \mathbf{v}_2 + \cdots + \frac{\mathbf{x} \cdot \mathbf{v}_m}{\mathbf{v}_m \cdot \mathbf{v}_m} \mathbf{v}_m. \quad (28.3)$$

Activity 28.2. Let $\mathbf{v}_1 = [1 \ 0 \ 1]^\top$, $\mathbf{v}_2 = [0 \ 1 \ 0]^\top$, and $\mathbf{v}_3 = [0 \ 0 \ 1]^\top$. The set $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is a basis for \mathbb{R}^3 . Let $\mathbf{x} = [1 \ 0 \ 0]^\top$. Calculate

$$\frac{\mathbf{x} \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1} \mathbf{v}_1 + \frac{\mathbf{x} \cdot \mathbf{v}_2}{\mathbf{v}_2 \cdot \mathbf{v}_2} \mathbf{v}_2 + \frac{\mathbf{x} \cdot \mathbf{v}_3}{\mathbf{v}_3 \cdot \mathbf{v}_3} \mathbf{v}_3.$$

Compare to \mathbf{x} . Does this violate Theorem 28.4? Explain.

Orthonormal Bases

The decomposition (28.3) is even simpler if $\mathbf{v}_k \cdot \mathbf{v}_k = 1$ for each k , that is, if \mathbf{v}_k is a unit vector for each k . In this case, the denominators are all 1 and we don't even need to consider them. We have a familiar example of such a basis for \mathbb{R}^n , namely the standard basis $\mathcal{S} = \{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$.

Recall that

$$\mathbf{v} \cdot \mathbf{v} = \|\mathbf{v}\|^2,$$

so the condition $\mathbf{v} \cdot \mathbf{v} = 1$ implies that the vector \mathbf{v} has norm 1. An orthogonal basis with this additional condition is a very nice basis and is given a special name.



Definition 28.5. An **orthonormal basis** $\mathcal{B} = \{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_m\}$ for a subspace W of \mathbb{R}^n is an orthogonal basis such that $\|\mathbf{u}_k\| = 1$ for $1 \leq k \leq m$.

In other words, an orthonormal basis is an orthogonal basis in which every basis vector is a unit vector. A good question to ask here is how we can construct an orthonormal basis from an orthogonal basis.

Activity 28.3.

- (a) Let \mathbf{v}_1 and \mathbf{v}_2 be orthogonal vectors. Explain how we can obtain unit vectors \mathbf{u}_1 in the direction of \mathbf{v}_1 and \mathbf{u}_2 in the direction of \mathbf{v}_2 .
- (b) Show that \mathbf{u}_1 and \mathbf{u}_2 from the previous part are orthogonal vectors.
- (c) Use the ideas from this problem to construct an orthonormal basis for \mathbb{R}^3 from the orthogonal basis $S = \left\{ \begin{bmatrix} -2 \\ 1 \\ -1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix} \right\}$.

In general, we can construct an orthonormal basis $\{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_m\}$ from an orthogonal basis $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_m\}$ by normalizing each vector in \mathcal{B} (that is, dividing each vector by its norm).

Orthogonal Matrices

We have seen in the diagonalization process that we diagonalize a matrix A with a matrix P whose columns are linearly independent eigenvectors of A . In general, calculating the inverse of the matrix whose columns are eigenvectors of A in the diagonalization process can be time consuming, but if the columns form an orthonormal set, then the calculation is very straightforward.

Activity 28.4. Let $\mathbf{u}_1 = \frac{1}{3}[2 \ 1 \ 2]^\top$, $\mathbf{u}_2 = \frac{1}{3}[-2 \ 2 \ 1]^\top$, and $\mathbf{u}_3 = \frac{1}{3}[1 \ 2 \ -2]^\top$. It is not difficult to see that the set $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ is an orthonormal basis for \mathbb{R}^3 . Let

$$A = [\mathbf{u}_1 \ \mathbf{u}_2 \ \mathbf{u}_3] = \frac{1}{3} \begin{bmatrix} 2 & -2 & 1 \\ 1 & 2 & 2 \\ 2 & 1 & -2 \end{bmatrix}.$$

- (a) Use the definition of the matrix-matrix product to find the entries of the second row of the matrix product $A^\top A$. Why should you have expected the result? (Hint: How are the rows of A^\top related to the columns of A ?)
- (b) With the result of part (a) in mind, what is the matrix product $A^\top A$? What does this tell us about the relationship between A^\top and A^{-1} ? Use technology to calculate A^{-1} and confirm your answer.
- (c) Suppose P is an $n \times n$ matrix whose columns form an orthonormal basis for \mathbb{R}^n . Explain why $P^\top P = I_n$.

The result of Activity 28.4 is that if the columns of a square matrix P form an orthonormal set, then $P^{-1} = P^T$. This makes calculating P^{-1} very easy. Note, however, that this only works if the columns of P form an orthonormal basis for $\text{Col } P$. You should also note that if P is an $n \times n$ matrix satisfying $P^T P = I_n$, then the columns of P must form an orthonormal set. Matrices like this appear quite often and are given a special name.

Definition 28.6. An **orthogonal** matrix is an $n \times n$ matrix P such that $P^T P = I_n$.¹

Activity 28.5. As a special case, we apply the result of Activity 28.4 to a 2×2 rotation matrix $P = \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix}$.

- (a) Show that the columns of P form an orthonormal set.
- (b) Use the fact that $P^{-1} = P^T$ to find P^{-1} . Explain how this shows that the inverse of a rotation matrix by an angle θ is just another rotation matrix but by the angle $-\theta$.

Orthogonal matrices are useful because they satisfy some special properties. For example, if P is an orthogonal $n \times n$ matrix and $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, then

$$(P\mathbf{x}) \cdot (P\mathbf{y}) = (P\mathbf{x})^T (P\mathbf{y}) = \mathbf{x}^T P^T P \mathbf{y} = \mathbf{x}^T \mathbf{y} = \mathbf{x} \cdot \mathbf{y}.$$

This property tells us that the linear transformation T defined by $T(\mathbf{x}) = P\mathbf{x}$ preserves dot products and, hence, orthogonality. In addition,

$$\|P\mathbf{x}\|^2 = P\mathbf{x} \cdot P\mathbf{x} = \mathbf{x} \cdot \mathbf{x} = \|\mathbf{x}\|^2,$$

so $\|P\mathbf{x}\| = \|\mathbf{x}\|$. This means that T preserves length. Such a transformation is called an *isometry* and it is convenient to work with functions that don't expand or contract things. Moreover, if \mathbf{x} and \mathbf{y} are nonzero vectors, then

$$\frac{P\mathbf{x} \cdot P\mathbf{y}}{\|P\mathbf{x}\| \|P\mathbf{y}\|} = \frac{\mathbf{x} \cdot \mathbf{y}}{\|\mathbf{x}\| \|\mathbf{y}\|}.$$

Thus T also preserves angles. Transformations defined by orthogonal matrices are very well behaved transformations. To summarize,

Theorem 28.7. Let P be an $n \times n$ orthogonal matrix and let $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$. Then

- (1) $(P\mathbf{x}) \cdot (P\mathbf{y}) = \mathbf{x} \cdot \mathbf{y}$,
- (2) $\|P\mathbf{x}\| = \|\mathbf{x}\|$, and
- (3) $\frac{P\mathbf{x} \cdot P\mathbf{y}}{\|P\mathbf{x}\| \|P\mathbf{y}\|} = \frac{\mathbf{x} \cdot \mathbf{y}}{\|\mathbf{x}\| \|\mathbf{y}\|}$ if \mathbf{x} and \mathbf{y} are nonzero.

We have discussed orthogonal and orthonormal bases for subspaces of \mathbb{R}^n in this section. There are several questions that follow, such as

- Can we always find an orthogonal (or orthonormal) basis for any subspace of \mathbb{R}^n ?
- Given a vector \mathbf{v} in W , can we find an orthogonal basis of W that contain \mathbf{v} ?
- Can we extend the concept of orthogonality to other vector spaces?

We will answer these questions in subsequent sections.

¹It isn't clear why such matrices are called orthogonal since the columns are actually orthonormal, but that is the standard terminology in mathematics.



Examples

What follows are worked examples that use the concepts from this section.

Example 28.8. Let $S = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$, where $\mathbf{v}_1 = [1 \ 1 \ -4]^T$, $\mathbf{v}_2 = [2 \ 2 \ 1]^T$, and $\mathbf{v}_3 = [1 \ -1 \ 0]^T$.

- (a) Show that S is an orthogonal set.
- (b) Create an orthonormal set $S' = \{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ from the vectors in S .
- (c) Just by calculating dot products, write the vector $\mathbf{w} = [2 \ 1 \ -1]^T$ as a linear combination of the vectors in S' .

Example Solution.

- (a) Using the dot product formula, we see that $\mathbf{v}_1 \cdot \mathbf{v}_2 = 0$, $\mathbf{v}_1 \cdot \mathbf{v}_3 = 0$, and $\mathbf{v}_2 \cdot \mathbf{v}_3 = 0$. Thus, the set S is an orthogonal set.
- (b) To make an orthonormal set $S' = \{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ from S , we divide each vector in S by its magnitude. This gives us

$$\mathbf{u}_1 = \frac{1}{\sqrt{18}}[1 \ 1 \ -4]^T, \quad \mathbf{u}_2 = \frac{1}{3}[2 \ 2 \ 1]^T, \quad \text{and} \quad \mathbf{u}_3 = \frac{1}{\sqrt{2}}[1 \ -1 \ 0]^T.$$

- (c) Since S' is an orthonormal basis for \mathbb{R}^3 , we know that

$$\begin{aligned} \mathbf{w} &= (\mathbf{w} \cdot \mathbf{u}_1)\mathbf{u}_1 + (\mathbf{w} \cdot \mathbf{u}_2)\mathbf{u}_2 + (\mathbf{w} \cdot \mathbf{u}_3)\mathbf{u}_3 \\ &= \frac{7}{\sqrt{18}}[1 \ 1 \ -4]^T + \frac{5}{3}[2 \ 2 \ 1]^T + \frac{1}{\sqrt{2}}[1 \ -1 \ 0]^T. \end{aligned}$$

Example 28.9. Let $\mathbf{u}_1 = \frac{1}{\sqrt{3}}[1 \ 1 \ 1]^T$, $\mathbf{u}_2 = \frac{1}{\sqrt{2}}[1 \ -1 \ 0]^T$, and $\mathbf{u}_3 = \frac{1}{\sqrt{6}}[1 \ 1 \ -2]^T$. Let $\mathcal{B} = \{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$.

- (a) Show that \mathcal{B} is an orthonormal basis for \mathbb{R}^3 .
- (b) Let $\mathbf{w} = [1 \ 2 \ 1]^T$. Find $[\mathbf{w}]_{\mathcal{B}}$.
- (c) Calculate $\|\mathbf{w}\|$ and $\|[\mathbf{w}]_{\mathcal{B}}\|$. What do you notice?
- (d) Show that the result of part (c) is true in general. That is, if $\mathcal{S} = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$ is an orthonormal basis for \mathbb{R}^n , and if $\mathbf{z} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \dots + c_n\mathbf{v}_n$, then

$$\|\mathbf{z}\| = \sqrt{c_1^2 + c_2^2 + \dots + c_n^2}.$$

Example Solution.

- (a) Using the dot product formula, we see that $\mathbf{u}_i \cdot \mathbf{u}_j = 0$ if $i \neq j$ and that $\mathbf{u}_i \cdot \mathbf{u}_i = 1$ for each i . Since orthogonal vectors are linearly independent, the set \mathcal{B} is a linearly independent set with 3 vectors in a 3-dimensional space. It follows that \mathcal{B} is an orthonormal basis for \mathbb{R}^3 .



(b) Since \mathcal{B} is an orthonormal basis for \mathbb{R}^3 , we know that

$$\mathbf{w} = (\mathbf{w} \cdot \mathbf{u}_1)\mathbf{u}_1 + (\mathbf{w} \cdot \mathbf{u}_2)\mathbf{u}_2 + (\mathbf{w} \cdot \mathbf{u}_3)\mathbf{u}_3.$$

Therefore,

$$[\mathbf{w}]_{\mathcal{B}} = [(\mathbf{w} \cdot \mathbf{u}_1) (\mathbf{w} \cdot \mathbf{u}_2) (\mathbf{w} \cdot \mathbf{u}_3)]^T = \left[\frac{4}{\sqrt{3}} - \frac{1}{\sqrt{2}} \frac{1}{\sqrt{6}} \right]^T.$$

(c) Using the definition of the norm of a vector we have

$$\begin{aligned} \|\mathbf{w}\| &= \sqrt{1^2 + 2^2 + 1^2} = \sqrt{6} \\ \|[\mathbf{w}]_{CB}\| &= \sqrt{\left(\frac{4}{\sqrt{3}}\right)^2 + \left(-\frac{1}{\sqrt{2}}\right)^2 + \left(\frac{1}{\sqrt{6}}\right)^2} = \sqrt{6}. \end{aligned}$$

So in this case we have $\|\mathbf{w}\| = \|[\mathbf{w}]_{\mathcal{B}}\|$.

(d) Let $\mathcal{S} = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$ be an orthonormal basis for \mathbb{R}^n , and suppose that $\mathbf{z} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \dots + c_n\mathbf{v}_n$. Then

$$\begin{aligned} \|\mathbf{z}\| &= \sqrt{\mathbf{z} \cdot \mathbf{z}} \\ &= \sqrt{(c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \dots + c_n\mathbf{v}_n) \cdot (c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \dots + c_n\mathbf{v}_n)}. \end{aligned} \quad (28.4)$$

Since \mathcal{S} is an orthonormal basis for \mathbb{R}^n , it follows that $\mathbf{v}_i \cdot \mathbf{v}_j = 0$ if $i \neq j$ and $\mathbf{v}_i \cdot \mathbf{v}_i = 1$. Expanding the dot product in (28.4), the only terms that won't be zero are the ones that involve $\mathbf{v}_i \cdot \mathbf{v}_i$. This leaves us with

$$\begin{aligned} \|\mathbf{z}\| &= \sqrt{(c_1\mathbf{v}_1 + \dots + c_n\mathbf{v}_n) \cdot (c_1\mathbf{v}_1 + \dots + c_n\mathbf{v}_n)} \\ &= \sqrt{c_1c_1(\mathbf{v}_1 \cdot \mathbf{v}_1) + c_2c_2(\mathbf{v}_2 \cdot \mathbf{v}_2) + \dots + c_nc_n(\mathbf{v}_n \cdot \mathbf{v}_n)} \\ &= \sqrt{c_1^2 + c_2^2 + \dots + c_n^2}. \end{aligned}$$

Summary

- A subset S of \mathbb{R}^n is an orthogonal set if $\mathbf{u} \cdot \mathbf{v} = 0$ for every pair of distinct vectors \mathbf{u} and \mathbf{v} in S .
- Any orthogonal set of nonzero vectors is linearly independent.
- A basis \mathcal{B} for a subspace W of \mathbb{R}^n is an orthogonal basis if \mathcal{B} is also an orthogonal set.
- An orthogonal basis \mathcal{B} for a subspace W of \mathbb{R}^n is an orthonormal basis if each vector in \mathcal{B} has unit length.
- If $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_m\}$ is an orthogonal basis for a subspace of \mathbb{R}^n and \mathbf{x} is any vector in W , then

$$\mathbf{x} = \sum_{i=1}^m c_i \mathbf{v}_i$$

where $c_i = \frac{\mathbf{x} \cdot \mathbf{v}_i}{\mathbf{v}_i \cdot \mathbf{v}_i}$.

- An $n \times n$ matrix P is an orthogonal matrix if $P^T P = I_n$. Orthogonal matrices are important, in part, because the matrix transformations they define are isometries.

Exercises

- (1) Find an orthogonal basis for the subspace $W = \{[x \ y \ z] : 4x - 3z = 0\}$ of \mathbb{R}^3 .
- (2) Let $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$ be an orthogonal basis for \mathbb{R}^n and, for some k between 1 and n , let $W = \text{Span}\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k\}$. Show that $\{\mathbf{v}_{k+1}, \mathbf{v}_{k+2}, \dots, \mathbf{v}_n\}$ is a basis for W^\perp .
- (3) Let W be a subspace of \mathbb{R}^n for some n , and let $\{\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_k\}$ be an orthogonal basis for W . Let \mathbf{x} be a vector in \mathbb{R}^n . and define \mathbf{w} as

$$\mathbf{w} = \frac{\mathbf{x} \cdot \mathbf{w}_1}{\mathbf{w}_1 \cdot \mathbf{w}_1} \mathbf{w}_1 + \frac{\mathbf{x} \cdot \mathbf{w}_2}{\mathbf{w}_2 \cdot \mathbf{w}_2} \mathbf{w}_2 + \cdots + \frac{\mathbf{x} \cdot \mathbf{w}_k}{\mathbf{w}_k \cdot \mathbf{w}_k} \mathbf{w}_k.$$

- (a) Explain why \mathbf{w} is in W .
- (b) Let $\mathbf{z} = \mathbf{x} - \mathbf{w}$. Show that \mathbf{z} is in W^\perp .
- (c) Explain why \mathbf{x} can be written as a sum of vectors, one in W and one in W^\perp .
- (d) Suppose $\mathbf{x} = \mathbf{w} + \mathbf{w}_1$ and $\mathbf{x} = \mathbf{u} + \mathbf{u}_1$, where \mathbf{w} and \mathbf{u} are in W and \mathbf{w}_1 and \mathbf{u}_1 are in W^\perp . Show that $\mathbf{w} = \mathbf{u}$ and $\mathbf{w}_1 = \mathbf{u}_1$, so that the representation of \mathbf{x} as a sum of a vector in W and a vector in W^\perp is unique.
- (4) Use the result of problem (3.) above and that $W \cap W^\perp = \{\mathbf{0}\}$ to show that $\dim(W) + \dim(W^\perp) = n$ for a subspace W of \mathbb{R}^n . (See Exercise 12 in Section 22 for the definition of the sum of subspaces.)
- (5) Let P be an $n \times n$ matrix. We showed that if P is an orthogonal matrix, then $(Px) \cdot (Py) = \mathbf{x} \cdot \mathbf{y}$ for any vectors \mathbf{x} and \mathbf{y} in \mathbb{R}^n . Now we ask if the converse of this statement is true. That is, determine the validity of the following statement: if $(Px) \cdot (Py) = \mathbf{x} \cdot \mathbf{y}$ for any vectors \mathbf{x} and \mathbf{y} in \mathbb{R}^n , then P is an orthogonal matrix? Verify your answer. (Hint: Consider $(Pe_i) \cdot (Pe_j)$ where e_t is the t th standard basis vector for \mathbb{R}^n .)
- (6) In this exercise we completely describe the 2×2 orthogonal matrices. Let P be an orthogonal 2×2 matrix.

- (a) The columns of P must be orthonormal vectors, so if we place the initial point of either of the columns of P at the origin, explain why its terminal point must have the form $(\cos(t), \sin(t))$ for some real number t .
- (b) As a consequence of part (a), let $[\cos(\theta) \ \sin(\theta)]^\top$ be the first column of P . Let the second column of P be $[\cos(\alpha) \ \sin(\alpha)]^\top$. Since the columns of P are orthogonal, how must the angle α be related to θ ?
- (c) Use a trigonometric identity to explain why

$$P = \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix}.$$

So the only orthogonal matrices in \mathbb{R}^2 are the rotation matrices.

- (7) Suppose A, B are orthogonal matrices of the same size.



- (a) Show that AB is also an orthogonal matrix.
- (b) Show that A^\top is also an orthogonal matrix.
- (c) Show that A^{-1} is also an orthogonal matrix.
- (8) Label each of the following statements as True or False. Provide justification for your response.
- (a) **True/False** Any orthogonal subset of \mathbb{R}^n is linearly independent.
- (b) **True/False** Every single vector set is an orthogonal set.
- (c) **True/False** If S is an orthogonal set in \mathbb{R}^n with exactly n nonzero vectors, then S is a basis for \mathbb{R}^n .
- (d) **True/False** Every set of three linearly independent vectors in \mathbb{R}^3 is an orthogonal set.
- (e) **True/False** If A and B are $n \times n$ orthogonal matrices, then $A + B$ must also be an orthogonal matrix.
- (f) **True/False** If the set $S = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$ is an orthogonal set in \mathbb{R}^n , then so is the set $\{c_1\mathbf{v}_1, c_2\mathbf{v}_2, \dots, c_n\mathbf{v}_n\}$ for any scalars c_1, c_2, \dots, c_n .
- (g) **True/False** If $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$ is an orthogonal basis of \mathbb{R}^n , then so is $\{c_1\mathbf{v}_1, c_2\mathbf{v}_2, \dots, c_n\mathbf{v}_n\}$ for any nonzero scalars c_1, c_2, \dots, c_n .
- (h) **True/False** If A is an $n \times n$ orthogonal matrix, the rows of A form an orthonormal basis of \mathbb{R}^n .
- (i) **True/False** If A is an orthogonal matrix, any matrix obtained by interchanging columns of A is also an orthogonal matrix.

Project: Understanding Rotations in 3-Space

Recall that a counterclockwise rotation of 2-space around the origin by an angle θ is accomplished by left multiplication by the matrix $\begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix}$. Notice that the columns of this rotation matrix are orthonormal, so this rotation matrix is an orthogonal matrix. As the next activity shows, rotation matrices in 3D are also orthogonal matrices.

Project Activity 28.1. Let R be a rotation matrix in 3D. A rotation does not change lengths of vectors, nor does it change angles between vectors. Let $\mathbf{e}_1 = [1 \ 0 \ 0]^\top$, $\mathbf{e}_2 = [0 \ 1 \ 0]^\top$, and $\mathbf{e}_3 = [0 \ 0 \ 1]^\top$ be the standard unit vectors in \mathbb{R}^3 .

- (a) Explain why the columns of R form an orthonormal set. (Hint: How are $R\mathbf{e}_1$, $R\mathbf{e}_2$, and $R\mathbf{e}_3$ related to the columns of R ?)
- (b) Explain why R is an orthogonal matrix. What must be true about $\det(R)$? (Hint: What is R^\top and what is $\det(R^\top R)$?)

By Project Activity 28.1 we know that the determinant of any rotation matrix is either 1 or -1 . Having a determinant of 1 preserves orientation, and we will identify these rotations as being



counterclockwise, and we will identify the others with determinant of -1 as being clockwise. We will set the convention that a rotation is always measured counterclockwise (as we did in \mathbb{R}^2), and so every rotation matrix will have determinant 1.

Returning to the counterclockwise rotation of 2-space around the origin by an angle θ determined by left multiplication by the matrix $\begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix}$, we can think of this rotation in 3-space as the rotation that keeps points in the xy plane in the xy plane, but rotates these points counterclockwise around the z axis. In other words, in the standard xyz coordinate system, with standard basis e_1, e_2, e_3 , our rotation matrix R has the property that $Re_3 = e_3$. Now Re_3 is the third column of R , so the third column of R is e_3 . Similarly, Re_1 is the first column of R and Re_2 is the second column of R . Since R is a counterclockwise rotation of the xy plane space around the origin by an angle θ it follows that this rotation is given by the matrix

$$R_{e_3}(\theta) = \begin{bmatrix} \cos(\theta) & -\sin(\theta) & 0 \\ \sin(\theta) & \cos(\theta) & 0 \\ 0 & 0 & 1 \end{bmatrix}. \quad (28.5)$$

In this notation in (28.5), the subscript gives the direction of the line fixed by the rotation and the angle provides the counterclockwise rotation in the plane perpendicular to this vector. This vector is called a *normal* vector for the rotation. Note also that the columns of $R_{e_3}(\theta)$ form an orthogonal set such that each column vector has norm 1.

This idea describes a general rotation matrix $R_n(\theta)$ in 3D by specifying a normal vector n and an angle θ . For example, with roll, a normal vector points from the tail of the aircraft to its tip. It is our goal to understand how we can determine an arbitrary rotation matrix of the form $R_n(\theta)$. We can accomplish this by using the rotation around the z axis and change of basis matrices to find rotation matrices around other axes. Let $S = \{e_1, e_2, e_3\}$ be the standard basis for \mathbb{R}^3 .

Project Activity 28.2. In this activity we see how to determine the rotation matrix around the x axis using the matrix $R_{e_3}(\theta)$ and a change of basis.

- (a) Define a new ordered basis B so that our axis of rotation is the third vector. So in this case the third vector in B will be e_1 . The other two vectors need to make B an orthonormal set. So we have plenty of choices. For example, we could set $B = \{e_2, e_3, e_1\}$. Find the change of basis matrix $P_{S \leftarrow B}$ from B to S .
- (b) Use the change of basis matrix from part (a) to find the change of basis matrix $P_{B \leftarrow S}$ from S to B .
- (c) To find our rotation matrix around the x axis, we can first change basis from S to B , then perform a rotation around the new z axis using (28.5), then changing basis back from B to S . In other words,

$$R_{e_1}(\theta) = P_{S \leftarrow B} R_{e_3}(\theta) P_{B \leftarrow S}.$$

Find the entries of this matrix $R_{e_1}(\theta)$.

IMPORTANT NOTE: We could have considered using $B_1 = \{e_3, e_2, e_1\}$ in Project Activity 28.2



instead of $\mathcal{B} = \{\mathbf{e}_2, \mathbf{e}_3, \mathbf{e}_1\}$. Then we would have

$$P_{S \leftarrow \mathcal{B}_1} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}.$$

The difference between the two options is that, in the first we have $\det(P_{S \leftarrow \mathcal{B}_1}) = -1$ while $\det(P_{S \leftarrow \mathcal{B}}) = 1$ in the second. Using \mathcal{B}_1 will give clockwise rotations while \mathcal{B} gives counterclockwise rotations (this is the difference between a left hand system and a right hand system). So it is important to ensure that our change of basis matrix is one with determinant 1.

Project Activity 28.3. Repeat Project Activity 28.3 to find the 3D rotation around the y axis.

We do one more example to illustrate the process before tackling the general case.

Project Activity 28.4. In this activity we find the rotation around the axis given by the line $x = y/2 = z$. This line is in the direction of the vector $\mathbf{n} = [1 2 1]^\top$. So we start with making a unit vector in the direction of \mathbf{n} as the third vector in an ordered basis \mathcal{B} . The other two vectors need to make \mathcal{B} an orthonormal set with $\det(P_{S \leftarrow \mathcal{B}}) = 1$.

- (a) Find a unit vector \mathbf{w} in the direction of \mathbf{n} .
- (b) Show that $[2 -1 0]^\top$ is orthogonal to the vector \mathbf{w} from part (a). Then find a unit vector \mathbf{v} that is in the same direction as $[2 -1 0]^\top$.
- (c) Let \mathbf{v} be as in the previous part. Now the trick is to find a third unit vector \mathbf{u} so that $\mathcal{B} = \{\mathbf{u}, \mathbf{v}, \mathbf{w}\}$ is an orthonormal set. This can be done with the cross product. If $\mathbf{a} = [a_1 a_2 a_3]^\top$ and $\mathbf{b} = [b_1 b_2 b_3]^\top$, then the cross product $\mathbf{a} \times \mathbf{b}$ of \mathbf{a} and \mathbf{b} is the vector

$$\mathbf{a} \times \mathbf{b} = (a_2 b_3 - a_3 b_2) \mathbf{e}_1 - (a_1 b_3 - a_3 b_1) \mathbf{e}_2 + (a_1 b_2 - a_2 b_1) \mathbf{e}_3.$$

(You can check that $\{\mathbf{a} \times \mathbf{b}, \mathbf{a}, \mathbf{b}\}$ is an orthogonal set that gives the correct determinant for the change of basis matrix.) Use the cross product to find a unit vector \mathbf{u} so that $\mathcal{B} = \{\mathbf{u}, \mathbf{v}, \mathbf{w}\}$ is an orthonormal set.

- (d) Find the entries of the matrix $R_{\mathbf{w}}(\theta)$.

In the next activity we summarize the general process to find a 3D rotation matrix $R_{\mathbf{n}}(\theta)$ for any normal vector \mathbf{n} . There is a GeoGebra applet at <https://www.geogebra.org/m/n9gbjhfx> that allows you to visualize rotation matrices in 3D.

Project Activity 28.5. Let $\mathbf{n} = [n_1 n_2 n_3]^\top$ be a normal vector (nonzero) for our rotation. We need to create an orthonormal basis $\mathcal{B} = \{\mathbf{u}, \mathbf{v}, \mathbf{w}\}$ where \mathbf{w} is a unit vector in the direction of \mathbf{n} so that the change of basis matrix $P_{S \leftarrow \mathcal{B}}$ has determinant 1.

- (a) Find, by inspection, a vector \mathbf{y} that is orthogonal to \mathbf{n} . (Hint: You may need to consider some cases to ensure that \mathbf{v} is not the zero vector.)



- (b) Once we have a normal vector \mathbf{n} and a vector \mathbf{y} orthogonal to \mathbf{n} , the vector $\mathbf{z} = \mathbf{y} \times \mathbf{n}$ gives us an orthogonal set $\{\mathbf{z}, \mathbf{y}, \mathbf{n}\}$. We then normalize each vector to create our orthonormal basis $\mathcal{B} = \{\mathbf{u}, \mathbf{v}, \mathbf{w}\}$. Use this process to find the matrix that produces a 45° counterclockwise rotation around the normal vector $[1 \ 0 \ -1]^\top$.

Section 29

Inner Products

Focus Questions

By the end of this section, you should be able to give precise and thorough answers to the questions listed below. You may want to keep these questions in mind to focus your thoughts as you complete the section.

- What is an inner product? What is an inner product space?
- What is an orthogonal set in an inner product space?
- What is an orthogonal basis for an inner product space?
- How do properties of orthogonality in R^n generalize to orthogonality in an inner product space?
- How do we find the coordinate vector for a vector in an inner product space relative to an orthogonal basis for the space?
- What is the projection of a vector orthogonal to a subspace and why are such orthogonal projections important?

Application: Fourier Series

In calculus, a Taylor polynomial for a function f is a polynomial approximation that fits f well around the center of the approximation. For this reason, Taylor polynomials are good *local* approximations, but they are not in general good global approximations. In particular, if a function f has periodic behavior it is impossible to model f well globally with polynomials that have infinite limits at infinity. For these kinds of functions, trigonometric polynomials are better choices. Trigonometric polynomials lead us to Fourier series, and we will investigate how inner products allow us to use trigonometric polynomials to model musical tones later in this section.

Introduction

We have seen that orthogonality in \mathbb{R}^n is an important concept. We can extend the idea of orthogonality, as well as the notions of length and angles, to a variety of different vector spaces beyond just \mathbb{R}^n as long as we have a product like a dot product. Such products are called *inner products*. Inner products lead us to many important ideas like Fourier series, wavelets, and others.

Recall that the dot product on \mathbb{R}^n assigns to each pair of vectors \mathbf{u} and \mathbf{v} the scalar $\mathbf{u} \cdot \mathbf{v}$. Thus, the dot product defines a mapping from $\mathbb{R}^n \times \mathbb{R}^n$ to \mathbb{R} . Recall also that the dot product is commutative, distributes over vector addition, and respects scalar multiplication. Additionally, the dot product of a vector by itself is always non-negative and is equal to 0 only when the vector is the zero vector. There is nothing special about using \mathbb{R}^n as the source for our vectors, and we can extend the notion of a dot product to any vector space using these properties.

Definition 29.1. An **inner product** $\langle \cdot, \cdot \rangle$ on a vector space V is a mapping from $V \times V \rightarrow \mathbb{R}$ satisfying

- (1) $\langle \mathbf{u}, \mathbf{v} \rangle = \langle \mathbf{v}, \mathbf{u} \rangle$ for all \mathbf{u} and \mathbf{v} in V ,
- (2) $\langle \mathbf{u} + \mathbf{v}, \mathbf{w} \rangle = \langle \mathbf{u}, \mathbf{w} \rangle + \langle \mathbf{v}, \mathbf{w} \rangle$ for all \mathbf{u} , \mathbf{v} , and \mathbf{w} in V ,
- (3) $\langle c\mathbf{u}, \mathbf{v} \rangle = c\langle \mathbf{u}, \mathbf{v} \rangle$ for all \mathbf{u} , \mathbf{v} in V and all scalars c ,
- (4) $\langle \mathbf{u}, \mathbf{u} \rangle \geq 0$ for all \mathbf{u} in V and $\langle \mathbf{u}, \mathbf{u} \rangle = 0$ if and only if $\mathbf{u} = \mathbf{0}$.

An **inner product space** is a vector space on which an inner product is defined.

Preview Activity 29.1.

- (1) Suppose we are given the mapping from $\mathbb{R}^2 \times \mathbb{R}^2$ to \mathbb{R} defined by

$$\langle \mathbf{u}, \mathbf{v} \rangle = u_1 v_1 + 2u_2 v_2$$

for $\mathbf{u} = [u_1 \ u_2]^\top$ and $\mathbf{v} = [v_1 \ v_2]^\top$ in \mathbb{R}^2 . Check that this mapping satisfies all of the properties of an inner product.

- (2)

- (3) Now consider the mapping from $\mathbb{R}^2 \times \mathbb{R}^2$ to \mathbb{R} defined by

$$\langle \mathbf{u}, \mathbf{v} \rangle = 2u_1 v_1 - 3u_2 v_2$$

for $\mathbf{u} = [u_1 \ u_2]^\top$ and $\mathbf{v} = [v_1 \ v_2]^\top$ in \mathbb{R}^2 . Show that this mapping does not satisfy the fourth property of an inner product.

- (4) Finally, show that the mapping from $\mathbb{P}_1 \times \mathbb{P}_1 \rightarrow \mathbb{R}$ defined by

$$\langle a_0 + a_1 t, b_0 + b_1 t \rangle = a_0 b_0 + a_1 b_1$$

for $a_0 + a_1 t, b_0 + b_1 t$ in \mathbb{P}_1 is an inner product on \mathbb{P}_1 .



Inner Product Spaces

As we saw in Definition 29.1, the idea of the dot product in \mathbb{R}^n can be extended to define inner products on different types of vector spaces. Preview Activity 29.1 provides two examples of inner products. The examples below provide some important inner products on vector spaces. Verification of the following as inner products is left to the exercises.

- If a_1, a_2, \dots, a_n are positive scalars, then

$$\langle [u_1 \ u_2 \ \cdots \ u_n]^T, [v_1 \ v_2 \ \cdots \ v_n]^T \rangle = a_1 u_1 v_1 + a_2 u_2 v_2 + \cdots + a_n u_n v_n$$

defines an inner product on \mathbb{R}^n .

- Every invertible $n \times n$ matrix A defines an inner product on \mathbb{R}^n by

$$\langle \mathbf{u}, \mathbf{v} \rangle = (A\mathbf{u}) \cdot (A\mathbf{v}).$$

- The definite integral defines an inner product:

$$\langle f, g \rangle = \int_a^b f(x)g(x) \, dx$$

for $f, g \in C[a, b]$ (where $C[a, b]$ is the vector space of all continuous functions on the interval $[a, b]$ – that $C[a, b]$ is a vector space is left for Exercise 1.)

- We can use the trace of a matrix (see Definition 18.7) to define an inner product on matrix spaces. If A and B are in the space $\mathcal{M}_{n \times n}$ of $n \times n$ matrices with real entries, we define the product $\langle A, B \rangle$ as

$$\langle A, B \rangle = \text{trace}(AB^T).$$

This defines an inner inner product on the space $\mathcal{M}_{n \times n}$ called the *Frobenius* inner product.

Since we defined inner products using the properties of the dot product, we might wonder if inner products actually satisfy all of the other properties of the dot product. For example, $\mathbf{u} \cdot \mathbf{0} = 0$ for any vector \mathbf{u} in \mathbb{R}^n ; but is it true that $\langle \mathbf{u}, \mathbf{0} \rangle = 0$ in every inner product space? This property is not part of the definition of an inner product, so we need to verify it if true.

Activity 29.1. Let \mathbf{u} be a vector in an inner product space V .

- Why is $\langle \mathbf{u}, \mathbf{0} \rangle = \langle \mathbf{u}, \mathbf{0} \rangle + \langle \mathbf{u}, \mathbf{0} \rangle$?
- How does the equation in part (a) show that $\langle \mathbf{u}, \mathbf{0} \rangle = 0$?

Activity 29.1 suggests that inner products share the defining properties of the dot product. Some properties of the inner product are given in the following theorem (the proofs of the remaining parts are left to the Exercises).

Theorem 29.2. Let $\langle \ , \ \rangle$ be an inner product on a vector space V and let \mathbf{u}, \mathbf{v} , and \mathbf{w} be vectors in V and c a scalar. Then



- (1) $\langle \mathbf{0}, \mathbf{v} \rangle = \langle \mathbf{v}, \mathbf{0} \rangle = 0$
- (2) $\langle \mathbf{u}, c\mathbf{v} \rangle = c\langle \mathbf{u}, \mathbf{v} \rangle$
- (3) $\langle \mathbf{u}, \mathbf{v} + \mathbf{w} \rangle = \langle \mathbf{u}, \mathbf{v} \rangle + \langle \mathbf{u}, \mathbf{w} \rangle$
- (4) $\langle \mathbf{u} - \mathbf{v}, \mathbf{w} \rangle = \langle \mathbf{u}, \mathbf{w} \rangle - \langle \mathbf{v}, \mathbf{w} \rangle$

Inner products will allow us to extend ideas of orthogonality, lengths of vectors, and angles between vectors to any inner product space.

The Length of a Vector

We can use inner products to define the length of any vector in an inner product space and the distance between two vectors in an inner product space. The idea comes right from the relationship between lengths of vectors in \mathbb{R}^n and the dot product (compare to Definition 27.2).

Definition 29.3. Let \mathbf{v} be a vector in an inner product space V . The **length** of \mathbf{v} is the real number

$$\|\mathbf{v}\| = \sqrt{\langle \mathbf{v}, \mathbf{v} \rangle}.$$

The length of a vector in a vector space is also called *magnitude* or *norm*. Just as in \mathbb{R}^n we can use the notion of length to define unit vectors in inner product spaces (compare to Definition 27.4).

Definition 29.4. A vector \mathbf{v} in inner product space is a **unit vector** if $\|\mathbf{v}\| = 1$.

We can find a unit vector in the direction of a nonzero vector \mathbf{v} in an inner product space V by dividing by the norm of the vector. That is, the vector $\frac{\mathbf{v}}{\|\mathbf{v}\|}$ is a unit vector in the direction of the vector \mathbf{v} , provided that \mathbf{v} is not zero.

We define the distance between vectors \mathbf{u} and \mathbf{v} in an inner product space V in the same way we defined distance in \mathbb{R}^n (compare to Definition 27.5).

Definition 29.5. Let \mathbf{u} and \mathbf{v} be vectors in an inner product space V . The **distance between \mathbf{u} and \mathbf{v}** is the length of the difference $\mathbf{u} - \mathbf{v}$ or

$$\|\mathbf{u} - \mathbf{v}\|.$$

Activity 29.2. Find the indicated length or distance in the inner product space.

- (a) Find the length of the vectors $\mathbf{u} = [1 \ 3]^\top$ and $\mathbf{v} = [3 \ 1]^\top$ using the inner product

$$\langle [u_1 \ u_2]^\top, [v_1 \ v_2]^\top \rangle = 2u_1v_1 + 3u_2v_2$$

in \mathbb{R}^2 .

- (b) Find the distance between the polynomials $p(t) = t + 1$ and $q(t) = t^2 - 1$ in $C[0, 1]$ using the inner product $\langle f, g \rangle = \int_0^1 f(x)g(x) dx$. (You may assume that \langle , \rangle defines an inner product on $C[0, 1]$, the space of continuous functions defined on the interval $[0, 1]$.)



Orthogonality in Inner Product Spaces

We defined orthogonality in \mathbb{R}^n using the dot product (see Definition 27.6) and the angle between vectors in \mathbb{R}^n . We can extend that idea to any inner product space.

We can define the angle between two vectors in an inner product space just as we did in \mathbb{R}^n . If \mathbf{u} and \mathbf{v} are nonzero vectors in an inner product space V , then the **angle θ between \mathbf{u} and \mathbf{v}** is such that

$$\cos(\theta) = \frac{|\langle \mathbf{u}, \mathbf{v} \rangle|}{\|\mathbf{u}\| \|\mathbf{v}\|}.$$

and $0 \leq \theta \leq \pi$. This angle is well-defined due to the Cauchy-Schwarz inequality $|\langle \mathbf{u}, \mathbf{v} \rangle| \leq \|\mathbf{u}\| \|\mathbf{v}\|$ whose proof is left to the exercises.

With the angle between vectors in mind, we can define orthogonal vectors in an inner product space.

Definition 29.6. Vectors \mathbf{u} and \mathbf{v} in an inner product space V are **orthogonal** if

$$\langle \mathbf{u}, \mathbf{v} \rangle = 0.$$

Note that this defines the zero vector to be orthogonal to every vector.

Activity 29.3.

- (a) Find a nonzero vector in \mathbb{R}^2 orthogonal to the vector $\mathbf{u} = [3 \ 1]^\top$ using the inner product $\langle [u_1 \ u_2]^\top, [v_1 \ v_2]^\top \rangle = 2u_1v_1 + 3u_2v_2$.

- (b) Determine if the vector $\mathbf{v} = \begin{bmatrix} 0 \\ 3 \\ -2 \end{bmatrix}$ is orthogonal to the vector $\mathbf{w} = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$ using the inner product $\langle \mathbf{u}, \mathbf{v} \rangle = (\mathbf{A}\mathbf{u}) \cdot (\mathbf{A}\mathbf{v})$ on \mathbb{R}^3 , where $\mathbf{A} = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 0 \end{bmatrix}$.

- (c) Find the angle between the two polynomials $p(t) = 1$ and $q(t) = t$ in \mathbb{P}_1 with inner product $\langle r(t), s(t) \rangle = \int_0^1 r(t)s(t) dt$.

Using orthogonality we can generalize the notions of orthogonal sets and bases, orthonormal bases and orthogonal complements we defined in \mathbb{R}^n to all inner product spaces in a natural way.

Orthogonal and Orthonormal Bases in Inner Product Spaces

As we did in \mathbb{R}^n , we define an orthogonal set to be one in which all of the vectors in the set are orthogonal to each other (compare to Definition 28.1).

Definition 29.7. A subset S of an inner product space V for which $\langle \mathbf{u}, \mathbf{v} \rangle = 0$ for all $\mathbf{u} \neq \mathbf{v}$ in S is called an **orthogonal set**.

As in \mathbb{R}^n , an orthogonal set of nonzero vectors is always linearly independent. The proof is similar to that of Theorem 28.3 and is left to the Exercises.



Theorem 29.8. Let $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_m\}$ be a set of nonzero orthogonal vectors in an inner product space V . Then the vectors $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_m$ are linearly independent.

A basis that is also an orthogonal set is given a special name (compare to Definition 28.2).

Definition 29.9. An **orthogonal basis** \mathcal{B} for a subspace W of an inner product space V is a basis of W that is also an orthogonal set.

Using the dot product in \mathbb{R}^n , we saw that the representation of a vector as a linear combination of vectors in an orthogonal or orthonormal basis was quite elegant. The same is true in any inner product space. To see this, let $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_m\}$ be an orthogonal basis for a subspace W of an inner product space V and let \mathbf{x} be any vector in W . We know that

$$\mathbf{x} = x_1 \mathbf{v}_1 + x_2 \mathbf{v}_2 + \cdots + x_m \mathbf{v}_m$$

for some scalars x_1, x_2, \dots, x_m . If $1 \leq k \leq m$, then, using inner product properties and the orthogonality of the vectors \mathbf{v}_i , we have

$$\langle \mathbf{v}_k, \mathbf{x} \rangle = x_1 \langle \mathbf{v}_k, \mathbf{v}_1 \rangle + x_2 \langle \mathbf{v}_k, \mathbf{v}_2 \rangle + \cdots + x_m \langle \mathbf{v}_k, \mathbf{v}_m \rangle = x_k \langle \mathbf{v}_k, \mathbf{v}_k \rangle.$$

So

$$x_k = \frac{\langle \mathbf{x}, \mathbf{v}_k \rangle}{\langle \mathbf{v}_k, \mathbf{v}_k \rangle}.$$

Thus, we can calculate each weight individually with two simple inner product calculations.

In other words, the coordinate vector $[\mathbf{x}]_{\mathcal{B}}$ of \mathbf{x} in an inner product space V with orthogonal basis $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_m\}$ is given by

$$[\mathbf{x}]_{\mathcal{B}} = \begin{bmatrix} \frac{\langle \mathbf{x}, \mathbf{v}_1 \rangle}{\langle \mathbf{v}_1, \mathbf{v}_1 \rangle} \\ \frac{\langle \mathbf{x}, \mathbf{v}_2 \rangle}{\langle \mathbf{v}_2, \mathbf{v}_2 \rangle} \\ \vdots \\ \frac{\langle \mathbf{x}, \mathbf{v}_m \rangle}{\langle \mathbf{v}_m, \mathbf{v}_m \rangle} \end{bmatrix}.$$

We summarize this discussion in the next theorem (compare to Theorem 28.4).

Theorem 29.10. Let $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_m\}$ be an orthogonal basis for a subspace of an inner product space V . Let \mathbf{x} be a vector in W . Then

$$\mathbf{x} = \frac{\langle \mathbf{x}, \mathbf{v}_1 \rangle}{\langle \mathbf{v}_1, \mathbf{v}_1 \rangle} \mathbf{v}_1 + \frac{\langle \mathbf{x}, \mathbf{v}_2 \rangle}{\langle \mathbf{v}_2, \mathbf{v}_2 \rangle} \mathbf{v}_2 + \cdots + \frac{\langle \mathbf{x}, \mathbf{v}_m \rangle}{\langle \mathbf{v}_m, \mathbf{v}_m \rangle} \mathbf{v}_m. \quad (29.1)$$

Activity 29.4. Let $p_1(t) = 1 - t$, $p_2(t) = -2 + 4t + 4t^2$, and $p_3(t) = 7 - 41t + 40t^2$ be vectors in the inner product space \mathbb{P}_2 with inner product defined by $\langle p(t), q(t) \rangle = \int_0^1 p(t)q(t) dt$. Let $\mathcal{B} = \{p_1(t), p_2(t), p_3(t)\}$. You may assume that \mathcal{B} is an orthogonal basis for \mathbb{P}_2 . Let $z(t) = 4 - 2t^2$. Find the weight x_3 so that $z(t) = x_1 p_1(t) + x_2 p_2(t) + x_3 p_3(t)$. Use technology as appropriate to evaluate any integrals.

The decomposition (29.1) is even simpler if $\langle \mathbf{v}_k, \mathbf{v}_k \rangle = 1$ for each k . Recall that

$$\langle \mathbf{v}, \mathbf{v} \rangle = \|\mathbf{v}\|^2,$$

so the condition $\langle \mathbf{v}, \mathbf{v} \rangle = 1$ implies that the vector \mathbf{v} has norm 1. As in \mathbb{R}^n , an orthogonal basis with this additional condition is given a special name (compare to Definition 28.5).



Definition 29.11. An **orthonormal basis** $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_m\}$ for a subspace W of an inner product space V is an orthogonal basis such that $\|\mathbf{v}_k\| = 1$ for $1 \leq k \leq m$.

If $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_m\}$ is an orthonormal basis for a subspace W of an inner product space V and \mathbf{x} is a vector in W , then (29.1) becomes

$$\mathbf{x} = \langle \mathbf{x}, \mathbf{v}_1 \rangle \mathbf{v}_1 + \langle \mathbf{x}, \mathbf{v}_2 \rangle \mathbf{v}_2 + \cdots + \langle \mathbf{x}, \mathbf{v}_m \rangle \mathbf{v}_m. \quad (29.2)$$

A good question to ask here is how we can construct an orthonormal basis from an orthogonal basis.

Activity 29.5. Consider vectors from an inner product space V .

- (a) Let \mathbf{v}_1 and \mathbf{v}_2 be orthogonal vectors. Explain how we can obtain unit vectors \mathbf{u}_1 in the direction of \mathbf{v}_1 and \mathbf{u}_2 in the direction of \mathbf{v}_2 .
- (b) Show that \mathbf{u}_1 and \mathbf{u}_2 from the previous part are orthogonal vectors.
- (c) Use the ideas from this problem to construct an orthonormal basis for the subspace

$$W = \text{Span} \left\{ \begin{bmatrix} 1 \\ 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \\ -1 \\ 2 \end{bmatrix}, \begin{bmatrix} 8 \\ 5 \\ -31 \\ -3 \end{bmatrix} \right\}$$

of the inner product space \mathbb{R}^4 with inner product

$$\langle [u_1 \ u_2 \ u_3 \ u_4]^T, [v_1 \ v_2 \ v_3 \ v_4]^T \rangle = 2u_1v_1 + 3u_2v_2 + u_3v_3 + 5u_4v_4.$$

(Note that you need to check for orthogonality.)

Orthogonal Projections onto Subspaces

Preview Activity 29.2. Let $\mathcal{B} = \{\mathbf{w}_1, \mathbf{w}_2\}$ be a basis for a subspace W of \mathbb{R}^3 , where $\mathbf{w}_1 = [1 \ 0 \ 0]^T$ and $\mathbf{w}_2 = [0 \ 1 \ 0]^T$. Note that \mathcal{B} is an orthonormal basis for W using the dot product as inner product. Let $\mathbf{v} = [1 \ 2 \ 1]^T$. Notice also that W is the xy -plane and that \mathbf{v} is not in W as illustrated in Figure 29.1.

- (1) Find the orthogonal projection \mathbf{u}_1 of \mathbf{v} onto $W_1 = \text{Span}\{\mathbf{w}_1\}$. (Hint: See Equation (27.2).)
- (2) Find the orthogonal projection \mathbf{u}_2 of \mathbf{v} onto $W_2 = \text{Span}\{\mathbf{w}_2\}$.
- (3) Calculate the distance between \mathbf{v} and \mathbf{u}_1 and \mathbf{v} and \mathbf{u}_2 . Which of \mathbf{u}_1 and \mathbf{u}_2 is closer to \mathbf{v} ?
- (4) Show that the vector $\frac{1}{2}[1 \ 4 \ 0]^T$ is in W and find the distance between \mathbf{v} and $\frac{1}{2}[1 \ 4 \ 0]^T$.
- (5) Part (4) shows that neither vector $\mathbf{u}_1 = \text{proj}_{\mathbf{w}_1} \mathbf{v}$ nor $\mathbf{u}_2 = \text{proj}_{\mathbf{w}_2} \mathbf{v}$ is the vector in W that is closest to \mathbf{v} . We should probably expect this since neither projection uses the fact that the other vector might contribute to the closest vector. Our goal is to find the linear combination \mathbf{w} of \mathbf{w}_1 and \mathbf{w}_2 in W that makes $\|\mathbf{w} - \mathbf{v}\|$ the smallest. Letting $\mathbf{w} = a\mathbf{w}_1 + b\mathbf{w}_2$, we have that $\|\mathbf{w} - \mathbf{v}\| = \sqrt{(a-1)^2 + (b-2)^2 + 1}$. Find the weights a and b that minimize this norm $\|\mathbf{w} - \mathbf{v}\|$.



- (6) A picture of \mathbf{w}_1 , \mathbf{w}_2 , W , and \mathbf{v} is shown in Figure 29.1. Draw in \mathbf{u}_1 , \mathbf{u}_2 , and the vector \mathbf{w} you found in part (5). There is a specific relationship between this vector \mathbf{w} and \mathbf{u}_1 and \mathbf{u}_2 . Describe this relationship algebraically and illustrate it graphically in Figure 29.1.

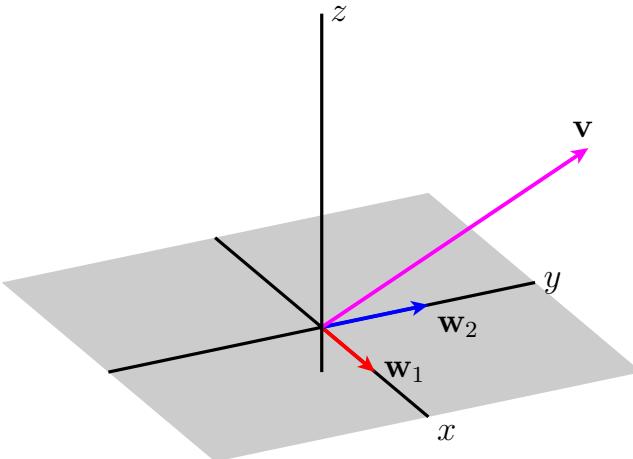


Figure 29.1: The space W and vectors \mathbf{w}_1 , \mathbf{w}_2 , and \mathbf{v}

Preview Activity 29.2 gives an indication of how we can project a vector \mathbf{v} in \mathbb{R}^n onto a subspace W of \mathbb{R}^n . If we have an orthogonal basis for W , we can just add the orthogonal projections of \mathbf{v} onto each basis vector. The resulting vector is called the orthogonal projection of \mathbf{v} onto the subspace W . As we did with orthogonal projections onto vectors, we can also define the projection of \mathbf{v} orthogonal to W . All of this can be done in the context of inner product spaces. Note that to make this all work out properly, we will need an orthogonal basis for W .

Definition 29.12. Let W be a subspace of an inner product space V and let $\mathcal{B} = \{\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_m\}$ be an orthogonal basis for W . For a vector \mathbf{v} in V , the **orthogonal projection of \mathbf{v} onto W** is the vector

$$\text{proj}_W \mathbf{v} = \frac{\langle \mathbf{v}, \mathbf{w}_1 \rangle}{\langle \mathbf{w}_1, \mathbf{w}_1 \rangle} \mathbf{w}_1 + \frac{\langle \mathbf{v}, \mathbf{w}_2 \rangle}{\langle \mathbf{w}_2, \mathbf{w}_2 \rangle} \mathbf{w}_2 + \cdots + \frac{\langle \mathbf{v}, \mathbf{w}_m \rangle}{\langle \mathbf{w}_m, \mathbf{w}_m \rangle} \mathbf{w}_m.$$

The **projection of \mathbf{v} orthogonal to W** is the vector

$$\text{proj}_{W^\perp} \mathbf{v} = \mathbf{v} - \text{proj}_W \mathbf{v}.$$

The notation $\text{proj}_{W^\perp} \mathbf{v}$ indicates that we expect this vector to be orthogonal to every vector in W .

Activity 29.6. Let $W = \text{Span}\{\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3\}$ in \mathbb{R}^4 , where $\mathbf{w}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 0 \end{bmatrix}$, $\mathbf{w}_2 = \begin{bmatrix} -1 \\ 1 \\ -1 \\ 2 \end{bmatrix}$, and

$\mathbf{w}_3 = \begin{bmatrix} 8 \\ 5 \\ -31 \\ -3 \end{bmatrix}$. Recall that we showed in Activity 29.5 that this was an orthogonal basis. Find

the projection of the vector $\mathbf{v} = \begin{bmatrix} 2 \\ 0 \\ 0 \\ 1 \end{bmatrix}$ onto W using the inner product

$$\langle [u_1 \ u_2 \ u_3 \ u_4]^T, [v_1 \ v_2 \ v_3 \ v_4]^T \rangle = 2u_1v_1 + 3u_2v_2 + u_3v_3 + 5u_4v_4.$$

Show directly that $\text{proj}_{W^\perp}\mathbf{v}$ is orthogonal to the basis vectors for W .

Activity 29.6 indicates that the vector $\text{proj}_{W^\perp}\mathbf{v}$ is in fact orthogonal to every vector in W . To see that this is true in general, let $\mathcal{B} = \{\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_m\}$ be an orthogonal basis for a subspace W of an inner product space V and let \mathbf{v} be a vector in V . Let

$$\mathbf{w} = \text{proj}_W\mathbf{v} = \frac{\langle \mathbf{v}, \mathbf{w}_1 \rangle}{\langle \mathbf{w}_1, \mathbf{w}_1 \rangle} \mathbf{w}_1 + \frac{\langle \mathbf{v}, \mathbf{w}_2 \rangle}{\langle \mathbf{w}_2, \mathbf{w}_2 \rangle} \mathbf{w}_2 + \cdots + \frac{\langle \mathbf{v}, \mathbf{w}_m \rangle}{\langle \mathbf{w}_m, \mathbf{w}_m \rangle} \mathbf{w}_m.$$

Then $\mathbf{v} - \mathbf{w}$ is the projection of \mathbf{v} orthogonal to W . We will show that $\mathbf{v} - \mathbf{w}$ is orthogonal to every basis vector for W . Since \mathcal{B} is an orthogonal basis for W , we know that $\mathbf{w}_i \cdot \mathbf{w}_j = 0$ for $i \neq j$. So if k is between 1 and m then

$$\begin{aligned} \langle \mathbf{w}_k, \mathbf{v} - \mathbf{w} \rangle &= \langle \mathbf{w}_k, \mathbf{v} \rangle - \langle \mathbf{w}_k, \mathbf{w} \rangle \\ &= \langle \mathbf{w}_k, \mathbf{v} \rangle - \left[\langle \mathbf{w}_k, \frac{\langle \mathbf{v}, \mathbf{w}_1 \rangle}{\langle \mathbf{w}_1, \mathbf{w}_1 \rangle} \mathbf{w}_1 + \cdots + \frac{\langle \mathbf{v}, \mathbf{w}_m \rangle}{\langle \mathbf{w}_m, \mathbf{w}_m \rangle} \mathbf{w}_m \rangle \right] \\ &= \langle \mathbf{w}_k, \mathbf{v} \rangle - \left(\frac{\langle \mathbf{v}, \mathbf{w}_k \rangle}{\langle \mathbf{w}_k, \mathbf{w}_k \rangle} \right) \langle \mathbf{w}_k, \mathbf{w}_k \rangle \\ &= \langle \mathbf{w}_k, \mathbf{v} \rangle - \langle \mathbf{v}, \mathbf{w}_k \rangle \\ &= 0. \end{aligned}$$

So the vector $\mathbf{v} - \mathbf{w}$ is orthogonal to every basis vector for W , and therefore to every vector in W (see Theorem 27.10). So, in fact, $\text{proj}_{W^\perp}\mathbf{v}$ is the projection of \mathbf{v} onto the orthogonal complement of W , which will be defined shortly.

Best Approximations

In many situations we are interested in approximating a vector that is not in a subspace with a vector in the subspace (e.g., linear regression to fit a line to a set of data). In these cases we usually want to find the vector in the subspace that “best” approximates the given vector using a specified inner product. As we will soon see, the projection of a vector onto a subspace has the property that the projection is the “best” approximation over all vectors in the subspace in terms of the length. In other words, $\text{proj}_W\mathbf{v}$ is the vector in W closest to \mathbf{v} and therefore the best approximation of \mathbf{v} by a vector in W . To see that this is true in any inner product space, we first need a generalization of the Pythagorean Theorem that holds in inner product spaces.

Theorem 29.13 (Generalized Pythagorean Theorem). *Let \mathbf{u} and \mathbf{v} be orthogonal vectors in an inner product space V . Then*

$$\|\mathbf{u} - \mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2.$$



Proof. Let \mathbf{u} and \mathbf{v} be orthogonal vectors in an inner product space V . Then

$$\begin{aligned} \|\mathbf{u} - \mathbf{v}\|^2 &= \langle \mathbf{u} - \mathbf{v}, \mathbf{u} - \mathbf{v} \rangle \\ &= \langle \mathbf{u}, \mathbf{u} \rangle - 2\langle \mathbf{u}, \mathbf{v} \rangle + \langle \mathbf{v}, \mathbf{v} \rangle \\ &= \langle \mathbf{u}, \mathbf{u} \rangle - 2(0) + \langle \mathbf{v}, \mathbf{v} \rangle \\ &= \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2. \end{aligned}$$

■

Note that replacing \mathbf{v} with $-\mathbf{v}$ in the theorem also shows that $\|\mathbf{u} + \mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2$ if \mathbf{u} and \mathbf{v} are orthogonal.

Now we will prove that the projection of a vector \mathbf{u} onto a subspace W of an inner product space V is the best approximation in W to the vector \mathbf{u} .

Theorem 29.14. *Let W be a subspace of an inner product space V and let \mathbf{u} be a vector in V . Then*

$$\|\mathbf{u} - \text{proj}_W \mathbf{u}\| < \|\mathbf{u} - \mathbf{x}\|$$

for every vector \mathbf{x} in W different from $\text{proj}_W \mathbf{u}$.

Proof. Let W be a subspace of an inner product space V and let \mathbf{u} be a vector in V . Let \mathbf{x} be a vector in W . Now

$$\mathbf{u} - \mathbf{x} = (\mathbf{u} - \text{proj}_W \mathbf{u}) + (\text{proj}_W \mathbf{u} - \mathbf{x}).$$

Since both $\text{proj}_W \mathbf{u}$ and \mathbf{x} are in W , we know that $\text{proj}_W \mathbf{u} - \mathbf{x}$ is in W . Since $\text{proj}_{W^\perp} \mathbf{u} = \mathbf{u} - \text{proj}_W \mathbf{u}$ is orthogonal to every vector in W , we have that $\mathbf{u} - \text{proj}_W \mathbf{u}$ is orthogonal to $\text{proj}_W \mathbf{u} - \mathbf{x}$. We can now use the Generalized Pythagorean Theorem to conclude that

$$\|\mathbf{u} - \mathbf{x}\|^2 = \|\mathbf{u} - \text{proj}_W \mathbf{u}\|^2 + \|\text{proj}_W \mathbf{u} - \mathbf{x}\|^2.$$

Since $\mathbf{x} \neq \text{proj}_W \mathbf{u}$, it follows that $\|\text{proj}_W \mathbf{u} - \mathbf{x}\|^2 > 0$ and

$$\|\mathbf{u} - \mathbf{x}\|^2 > \|\mathbf{u} - \text{proj}_W \mathbf{u}\|^2.$$

Since norms are nonnegative, we can conclude that $\|\mathbf{u} - \text{proj}_W \mathbf{u}\| < \|\mathbf{u} - \mathbf{x}\|$ as desired. ■

Theorem 29.14 shows that the distance from $\text{proj}_W \mathbf{v}$ to \mathbf{v} is less than the distance from any other vector in W to \mathbf{v} . So $\text{proj}_W \mathbf{v}$ is the best approximation to \mathbf{v} of all the vectors in W .

In \mathbb{R}^n using the dot product as inner product, if $\mathbf{v} = [v_1 \ v_2 \ v_3 \ \dots \ v_n]^\top$ and $\text{proj}_W \mathbf{v} = [w_1 \ w_2 \ w_3 \ \dots \ w_n]^\top$, then the square of the error in approximating \mathbf{v} by $\text{proj}_W \mathbf{v}$ is given by

$$\|\mathbf{v} - \text{proj}_W \mathbf{v}\|^2 = \sum_{i=1}^n (v_i - w_i)^2.$$

So $\text{proj}_W \mathbf{v}$ minimizes this sum of squares over all vectors in W . As a result, we call $\text{proj}_W \mathbf{v}$ the *least squares approximation* to \mathbf{v} .

Activity 29.7. The set $\mathcal{B} = \{1, t - \frac{1}{2}, t^3 - \frac{9}{10}t + \frac{1}{5}\}$ is an orthogonal basis for a subspace W of the inner product space \mathbb{P}_3 using the inner product $\langle p(t), q(t) \rangle = \int_0^1 p(t)q(t) dt$. Find the polynomial in W that is closest to the polynomial $r(t) = t^2$ and give a numeric estimate of how good this approximation is.



Orthogonal Complements

If we have a set of vectors S in an inner product space V , we can define the orthogonal complement of S as we did in \mathbb{R}^n (see 27.8).

Definition 29.15. The **orthogonal complement** of a subset S of an inner product space V is the set

$$S^\perp = \{\mathbf{v} \in V : \langle \mathbf{v}, \mathbf{u} \rangle = 0 \text{ for all } \mathbf{u} \in S\}.$$

As we saw in \mathbb{R}^n , to show that a vector is in the orthogonal complement of a subspace, it is enough to show that the vector is orthogonal to every vector in a basis for that subspace (see Theorem 27.10). The proof is left for the exercises.

Theorem 29.16. Let $\mathcal{B} = \{\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_m\}$ be a basis for a subspace W of an inner product space V . A vector \mathbf{v} in V is orthogonal to every vector in W if and only if \mathbf{v} is orthogonal to every vector in \mathcal{B} .

Activity 29.8. Consider \mathbb{P}_2 with the inner product $\langle p(t), q(t) \rangle = \int_0^1 p(t)q(t) dt$.

- (a) Find $\langle p(t), 1 - t \rangle$ where $p(t) = a + bt + ct^2$ is in \mathbb{P}_2 .
- (b) Describe as best you can the orthogonal complement of $\text{Span}\{1 - t\}$ in \mathbb{P}_2 . Is $p(t) = 1 - 2t - 2t^2$ in this orthogonal complement? Is $p(t) = 1 + t - t^2$?

To conclude this section we will investigate an important connection between a subspace W and W^\perp .

Activity 29.9. Let V be an inner product space of dimension n , and let W be a subspace of V . Let \mathbf{x} be any vector in V . We will demonstrate that \mathbf{x} can be written uniquely as a sum of a vector in W and a vector in W^\perp .

- (a) Explain why $\text{proj}_W \mathbf{x}$ is in W .
- (b) Explain why $\text{proj}_{W^\perp} \mathbf{x}$ is in W^\perp .
- (c) Explain why \mathbf{x} can be written as a sum of vectors, one in W and one in W^\perp .
- (d) Now we demonstrate the uniqueness of this decomposition. Suppose $\mathbf{x} = \mathbf{w} + \mathbf{w}_1$ and $\mathbf{x} = \mathbf{u} + \mathbf{u}_1$, where \mathbf{w} and \mathbf{u} are in W and \mathbf{w}_1 and \mathbf{u}_1 are in W^\perp . Show that $\mathbf{w} = \mathbf{u}$ and $\mathbf{w}_1 = \mathbf{u}_1$, so that the representation of \mathbf{x} as a sum of a vector in W and a vector in W^\perp is unique. (Hint: What is $W \cap W^\perp$?)

We summarize the result of Activity 29.9.

Theorem 29.17. Let V be a finite dimensional inner product space, and let W be a subspace of V . Any vector in V can be written in a unique way as a sum of a vector in W and a vector in W^\perp .

Theorem 29.17 is useful in many applications. For example, to compress an image using wavelets, we store the image as a collection of data, then rewrite the data using a succession of subspaces and their orthogonal complements. This new representation allows us to visualize the data in a way that compression is possible.



Examples

What follows are worked examples that use the concepts from this section.

Example 29.18. Let $V = \mathbb{P}_3$ be the inner product space with inner product

$$\langle p(t), q(t) \rangle = \int_{-1}^1 p(t)q(t) dt.$$

Let $p_1(t) = 1 + t$, $p_2(t) = 1 - 3t$, $p_3(t) = 3t - 5t^3$, and $p_4(t) = 1 - 3t^2$.

- (a) Show that the set $\mathcal{B} = \{p_1(t), p_2(t), p_3(t), p_4(t)\}$ is an orthogonal basis for V .
- (b) Use 29.10 to write the polynomial $q(t) = t^2 + t^3$ as a linear combination of the basis vectors in \mathcal{B} .

Example Solution. All calculations are done by hand or with a computer algebra system, so we leave those details to the reader.

- (a) If we show that the set \mathcal{B} is an orthogonal set, then Theorem 29.8 shows that \mathcal{B} is linearly independent. Since $\dim(\mathbb{P}_3) = 4$, the linearly independent set \mathcal{B} that contains four vectors must be a basis for \mathbb{P}_3 .

To determine if the set \mathcal{B} is an orthogonal set, we must calculate the inner products of pairs of distinct vectors in \mathcal{B} . Since $\langle 1+t, 1-3t \rangle = 0$, $\langle 1+t, 3t-5t^3 \rangle = 0$, $\langle 1+t, 1-3t^2 \rangle = 0$, $\langle 1-3t, 3t-5t^3 \rangle = 0$, $\langle 1-3t, 1-3t^2 \rangle = 0$, and $\langle 3t-5t^3, 1-3t^2 \rangle = 0$, we conclude that \mathcal{B} is an orthogonal basis for \mathbb{P}_3 .

- (b) We can write the polynomial $q(t) = 1 + t + t^2 + t^3$ as a linear combination of the basis vectors in \mathcal{B} as follows:

$$\begin{aligned} q(t) &= \frac{\langle q(t), p_1(t) \rangle}{\langle p_1(t), p_1(t) \rangle} p_1(t) + \frac{\langle q(t), p_2(t) \rangle}{\langle p_2(t), p_2(t) \rangle} p_2(t) \\ &\quad + \frac{\langle q(t), p_3(t) \rangle}{\langle p_3(t), p_3(t) \rangle} p_3(t) + \frac{\langle q(t), p_4(t) \rangle}{\langle p_4(t), p_4(t) \rangle} p_4(t). \end{aligned}$$

Now

$$\begin{aligned} \langle q(t), p_1(t) \rangle &= \frac{16}{15}, & \langle q(t), p_2(t) \rangle &= -\frac{8}{15}, & \langle q(t), p_3(t) \rangle &= -\frac{8}{35}, \\ \langle q(t), p_4(t) \rangle &= -\frac{8}{15}, & \langle p_1(t), p_1(t) \rangle &= \frac{8}{3}, & \langle p_2(t), p_2(t) \rangle &= 8, \\ \langle p_3(t), p_3(t) \rangle &= \frac{8}{7}, & \langle p_4(t), p_4(t) \rangle &= \frac{8}{5} \end{aligned}$$

so

$$\begin{aligned} q(t) &= \frac{\frac{16}{15}}{\frac{8}{3}} p_1(t) - \frac{\frac{8}{15}}{8} p_2(t) - \frac{\frac{8}{35}}{\frac{8}{7}} p_3(t) - \frac{\frac{8}{15}}{\frac{8}{5}} p_4(t) \\ &= \frac{2}{5} p_1(t) - \frac{1}{15} p_2(t) - \frac{1}{5} p_3(t) - \frac{1}{3} p_4(t). \end{aligned}$$

Example 29.19. Let V be the inner product space \mathbb{R}^4 with inner product defined by

$$\langle [u_1 \ u_2 \ u_3 \ u_4]^T, [v_1 \ v_2 \ v_3 \ v_4]^T \rangle = u_1 v_1 + 2u_2 v_2 + 3u_3 v_3 + 4u_4 v_4.$$



- (a) Let W be the plane spanned by $[-1 \ 1 \ 0 \ 1]^\top$ and $[6 \ 1 \ 7 \ 1]^\top$ in V . Find the vector in W that is closest to the vector $[2 \ 0 \ -1 \ 3]^\top$. Exactly how close is your best approximation to the vector $[2 \ 0 \ -1 \ 3]^\top$?
- (b) Express the vector $[2 \ 0 \ -1 \ 3]^\top$ as the sum of a vector in W and a vector orthogonal to W .

Example Solution.

- (a) The vector we're looking for is the projection of $[2 \ 0 \ -1 \ 3]^\top$ onto the plane. A spanning set for the plane is $\mathcal{B} = \{[-1 \ 1 \ 0 \ 1]^\top, [6 \ 1 \ 7 \ 1]^\top\}$. Neither vector in \mathcal{B} is a scalar multiple of the other, so \mathcal{B} is a basis for the plane. Since

$$\langle [-1 \ 1 \ 0 \ 1]^\top, [6 \ 1 \ 7 \ 1]^\top \rangle = -6 + 2 + 0 + 4 = 0,$$

the set \mathcal{B} is an orthogonal basis for the plane.

The projection of the vector $\mathbf{v} = [2 \ 0 \ -1 \ 3]^\top$ onto the plane spanned by $\mathbf{w}_1 = [-1 \ 1 \ 0 \ 1]^\top$ and $\mathbf{w}_2 = [6 \ 1 \ 7 \ 1]^\top$ is given by

$$\begin{aligned} \frac{\langle \mathbf{v}, \mathbf{w}_1 \rangle}{\langle \mathbf{w}_1, \mathbf{w}_1 \rangle} \mathbf{w}_1 + \frac{\langle \mathbf{v}, \mathbf{w}_2 \rangle}{\langle \mathbf{w}_2, \mathbf{w}_2 \rangle} \mathbf{w}_2 &= \frac{10}{7} [-1 \ 1 \ 0 \ 1]^\top + \frac{3}{189} [6 \ 1 \ 7 \ 1]^\top \\ &= \frac{1}{189} [-252 \ 273 \ 21 \ 273]^\top \\ &= \left[-\frac{4}{3} \ \frac{13}{9} \ \frac{1}{9} \ \frac{13}{9} \right]. \end{aligned}$$

To measure how close $\left[-\frac{4}{3} \ \frac{13}{9} \ \frac{1}{9} \ \frac{13}{9} \right]$ is to $[2 \ 0 \ -1 \ 3]^\top$, we calculate

$$\begin{aligned} \left\| \left[-\frac{4}{3} \ \frac{13}{9} \ \frac{1}{9} \ \frac{13}{9} \right] - [2 \ 0 \ -1 \ 3]^\top \right\| &= \left\| \left[-\frac{10}{3} \ \frac{13}{9} \ \frac{10}{9} \ -\frac{14}{9} \right] \right\| \\ &= \sqrt{\frac{100}{9} + \frac{338}{81} + \frac{300}{81} + \frac{784}{81}} \\ &= \frac{1}{9} \sqrt{2322} \\ &\approx 5.35. \end{aligned}$$

- (b) If $\mathbf{v} = [2 \ 0 \ -1 \ 3]^\top$, then $\text{proj}_W \mathbf{v}$ is in W and

$$\text{proj}_{W^\perp} \mathbf{v} = \mathbf{v} - \text{proj}_W \mathbf{v} = \left[-\frac{10}{3} \ \frac{13}{9} \ \frac{10}{9} \ -\frac{14}{9} \right]$$

is in W^\perp , and $\mathbf{v} = \text{proj}_W \mathbf{v} + \text{proj}_{W^\perp} \mathbf{v}$.

Summary

- An inner product $\langle \cdot, \cdot \rangle$ on a vector space V is a mapping from $V \times V \rightarrow \mathbb{R}$ satisfying
 - $\langle \mathbf{u}, \mathbf{v} \rangle = \langle \mathbf{v}, \mathbf{u} \rangle$ for all \mathbf{u} and \mathbf{v} in V ,



- (2) $\langle \mathbf{u} + \mathbf{v}, \mathbf{w} \rangle = \langle \mathbf{u}, \mathbf{w} \rangle + \langle \mathbf{v}, \mathbf{w} \rangle$ for all \mathbf{u} , \mathbf{v} , and \mathbf{w} in V ,
- (3) $\langle c\mathbf{u}, \mathbf{v} \rangle = c\langle \mathbf{u}, \mathbf{v} \rangle$ for all \mathbf{u} , \mathbf{v} in V and $c \in \mathbb{R}$,
- (4) $\langle \mathbf{u}, \mathbf{u} \rangle \geq 0$ for all \mathbf{u} in V and $\langle \mathbf{u}, \mathbf{u} \rangle = 0$ if and only if $\mathbf{u} = \mathbf{0}$.

- An inner product space is a pair $V, \langle \cdot, \cdot \rangle$ where V is a vector space and $\langle \cdot, \cdot \rangle$ is an inner product on V .
- The length of a vector \mathbf{v} in an inner product space V is defined to be the real number $\|\mathbf{v}\| = \sqrt{\langle \mathbf{v}, \mathbf{v} \rangle}$.
- The distance between two vectors \mathbf{u} and \mathbf{v} in an inner product space V is the scalar $\|\mathbf{u} - \mathbf{v}\|$.
- The angle θ between two vectors \mathbf{u} and \mathbf{v} is the angle which satisfies $0 \leq \theta \leq \pi$ and

$$\cos(\theta) = \frac{\langle \mathbf{u}, \mathbf{v} \rangle}{\|\mathbf{u}\| \|\mathbf{v}\|}.$$

- Two vectors \mathbf{u} and \mathbf{v} in an inner product space V are orthogonal if $\langle \mathbf{u}, \mathbf{v} \rangle = 0$.
- A subset S of an inner product space is an orthogonal set if $\langle \mathbf{u}, \mathbf{v} \rangle = 0$ for all $\mathbf{u} \neq \mathbf{v}$ in S .
- A basis for a subspace of an inner product space is an orthogonal basis if the basis is also an orthogonal set.
- Let $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_m\}$ be an orthogonal basis for a subspace of an inner product space V . Let \mathbf{x} be a vector in V . Then

$$\mathbf{x} = \sum_{i=1}^m c_i \mathbf{v}_i,$$

where

$$c_i = \frac{\langle \mathbf{x}, \mathbf{v}_i \rangle}{\langle \mathbf{v}_i, \mathbf{v}_i \rangle}$$

for each i .

- An orthogonal basis $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_m\}$ for a subspace W of an inner product space V is an orthonormal basis if $\|\mathbf{v}_k\| = 1$ for each k from 1 to m .
- If $\mathcal{B} = \{\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_m\}$ is an orthogonal basis for V and $\mathbf{x} \in V$, then

$$[\mathbf{x}]_{\mathcal{B}} = \begin{bmatrix} \frac{\langle \mathbf{x}, \mathbf{w}_1 \rangle}{\langle \mathbf{w}_1, \mathbf{w}_1 \rangle} \\ \frac{\langle \mathbf{x}, \mathbf{w}_2 \rangle}{\langle \mathbf{w}_2, \mathbf{w}_2 \rangle} \\ \vdots \\ \frac{\langle \mathbf{x}, \mathbf{w}_m \rangle}{\langle \mathbf{w}_m, \mathbf{w}_m \rangle} \end{bmatrix}.$$

- The projection of the vector \mathbf{v} in an inner product space V onto a subspace W of V is the vector

$$\text{proj}_W \mathbf{v} = \frac{\langle \mathbf{v}, \mathbf{w}_1 \rangle}{\langle \mathbf{w}_1, \mathbf{w}_1 \rangle} \mathbf{w}_1 + \frac{\langle \mathbf{v}, \mathbf{w}_2 \rangle}{\langle \mathbf{w}_2, \mathbf{w}_2 \rangle} \mathbf{w}_2 + \cdots + \frac{\langle \mathbf{v}, \mathbf{w}_m \rangle}{\langle \mathbf{w}_m, \mathbf{w}_m \rangle} \mathbf{w}_m,$$

where $\{\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_m\}$ is an orthogonal basis of W . Projections are important in that $\text{proj}_W \mathbf{v}$ is the best approximation of the vector \mathbf{v} by a vector in W in the least squares sense.



- With W as in (a), the projection of \mathbf{v} orthogonal to W is the vector

$$\text{proj}_{W^\perp} \mathbf{v} = \mathbf{v} - \text{proj}_W \mathbf{v}.$$

The norm of $\text{proj}_{W^\perp} \mathbf{v}$ provides a measure of how well $\text{proj}_W \mathbf{v}$ approximates the vector \mathbf{v} .

- The orthogonal complement of a subset S of an inner product space V is the set

$$S^\perp = \{\mathbf{v} \in V : \langle \mathbf{v}, \mathbf{u} \rangle = 0 \text{ for all } \mathbf{u} \in S\}.$$

Exercises

- (1) Let $C[a, b]$ be the set of all continuous real valued functions on the interval $[a, b]$. If f is in $C[a, b]$, we can extend f to a continuous function from \mathbb{R} to \mathbb{R} by letting F be the function defined by

$$F(x) = \begin{cases} f(a) & \text{if } x < a \\ f(x) & \text{if } a \leq x \leq b \\ f(b) & \text{if } b < x \end{cases}$$

In this way we can view $C[a, b]$ as a subset of \mathcal{F} , the vector space of all functions from \mathbb{R} to \mathbb{R} . Verify that $C[a, b]$ is a vector space.

- (2) Use the definition of an inner product to determine which of the following defines an inner product on the indicated space. Verify your answers.

- $\langle \mathbf{u}, \mathbf{v} \rangle = u_1 v_1 - u_2 v_1 - u_1 v_2 + 3u_2 v_2$ for $\mathbf{u} = [u_1 \ u_2]^\top$ and $\mathbf{v} = [v_1 \ v_2]^\top$ in \mathbb{R}^2
- $\langle f, g \rangle = \int_a^b f(x)g(x) dx$ for $f, g \in C[a, b]$ (where $C[a, b]$ is the vector space of all continuous functions on the interval $[a, b]$)
- $\langle f, g \rangle = f'(0)g'(0)$ for $f, g \in D(-1, 1)$ (where $D(a, b)$ is the vector space of all differentiable functions on the interval (a, b))
- $\langle \mathbf{u}, \mathbf{v} \rangle = (\mathbf{A}\mathbf{u}) \cdot (\mathbf{A}\mathbf{v})$ for $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$ and A an invertible $n \times n$ matrix

- (3) We can sometimes visualize an inner product in \mathbb{R}^2 or \mathbb{R}^3 (or other spaces) by describing the unit circle S^1 , where

$$S^1 = \{\mathbf{v} \in V : \|\mathbf{v}\| = 1\}$$

in that inner product space. For example, in the inner product space \mathbb{R}^2 with the dot product as inner product, the unit circle is just our standard unit circle. Inner products, however, can distort this familiar picture of the unit circle. Describe the points on the unit circle S^1 in the inner product space \mathbb{R}^2 with inner product $\langle [u_1 \ u_2], [v_1 \ v_2] \rangle = 2u_1 v_1 + 3u_2 v_2$ using the following steps.

- Let $\mathbf{x} = [x \ y] \in \mathbb{R}^2$. Set up an equation in x and y that is equivalent to the vector equation $\|\mathbf{x}\| = 1$.
- Describe the graph of the equation you found in \mathbb{R}^2 . It should have a familiar form. Draw a picture to illustrate. What do you think of calling this graph a “circle”?



- (4) Define $\langle \cdot, \cdot \rangle$ on \mathbb{R}^2 by $\langle [u_1 \ u_2]^T, [v_1 \ v_2]^T \rangle = 4u_1v_1 + 2u_2v_2$.
- Show that $\langle \cdot, \cdot \rangle$ is an inner product.
 - The inner product $\langle \cdot, \cdot \rangle$ can be represented as a matrix transformation $\langle \mathbf{u}, \mathbf{v} \rangle = \mathbf{u}^T A \mathbf{v}$, where \mathbf{u} and \mathbf{v} are written as column vectors. Find a matrix A that represents this inner product.
- (5) This exercise is a generalization of Exercise 4. Define $\langle \cdot, \cdot \rangle$ on \mathbb{R}^n by
- $$\langle [u_1 \ u_2 \ \cdots \ u_n]^T, [v_1 \ v_2 \ \cdots \ v_n]^T \rangle = a_1u_1v_1 + a_2u_2v_2 + \cdots + a_nu_nv_n$$
- for some positive scalars a_1, a_2, \dots, a_n .
- Show that $\langle \cdot, \cdot \rangle$ is an inner product.
 - The inner product $\langle \cdot, \cdot \rangle$ can be represented as a matrix transformation $\langle \mathbf{u}, \mathbf{v} \rangle = \mathbf{u}^T A \mathbf{v}$, where \mathbf{u} and \mathbf{v} are written as column vectors. Find a matrix A that represents this inner product.
- (6) Is the sum of two inner products on an inner product space V an inner product on V ? If yes, prove it. If no, provide a counterexample. (By the sum of inner products we mean a function $\langle \cdot, \cdot \rangle$ satisfying
- $$\langle \mathbf{u}, \mathbf{v} \rangle = \langle \mathbf{u}, \mathbf{v} \rangle_1 + \langle \mathbf{u}, \mathbf{v} \rangle_2$$
- for all \mathbf{u} and \mathbf{v} in V , where $\langle \cdot, \cdot \rangle_1$ and $\langle \cdot, \cdot \rangle_2$ are inner products on V .)
- (7) (a) Does $\langle \mathbf{u}, \mathbf{v} \rangle = \mathbf{u}^T A \mathbf{v}$ define an inner product on \mathbb{R}^n for every $n \times n$ matrix A ? Verify your answer.
- (b) If your answer to part (a) is no, are there any types of matrices for which $\langle \mathbf{u}, \mathbf{v} \rangle = \mathbf{u}^T A \mathbf{v}$ defines an inner product? (Hint: See Exercises 4 and 5.)
- (8) The trace of an $n \times n$ matrix $A = [a_{ij}]$ has some useful properties.
- Show that $\text{trace}(A + B) = \text{trace}(A) + \text{trace}(B)$ for any $n \times n$ matrices A and B .
 - Show that $\text{trace}(cA) = c\text{trace}(A)$ for any $n \times n$ matrix A and any scalar c .
 - Show that $\text{trace}(A^T) = \text{trace}(A)$ for any $n \times n$ matrix.
- (9) Let V be an inner product space and \mathbf{u}, \mathbf{v} be two vectors in V .
- Check that if $\mathbf{v} = \mathbf{0}$, the Cauchy-Schwarz inequality
- $$|\langle \mathbf{u}, \mathbf{v} \rangle| \leq \|\mathbf{u}\| \|\mathbf{v}\|$$
- holds.
- Assume $\mathbf{v} \neq \mathbf{0}$. Let $\lambda = \langle \mathbf{u}, \mathbf{v} \rangle / \|\mathbf{v}\|^2$ and $\mathbf{w} = \mathbf{u} - \lambda \mathbf{v}$. Use the fact that $\|\mathbf{w}\|^2 \geq 0$ to conclude the Cauchy-Schwarz inequality in this case.
- (10) The *Frobenius inner product* is defined as
- $$\langle A, B \rangle = \text{trace}(AB^T).$$

for $n \times n$ matrices A and B . Verify that $\langle A, B \rangle$ defines an inner product on $\mathcal{M}_{n \times n}$.



(11) Let $A = [a_{ij}]$ and $B = [b_{ij}]$ be two $n \times n$ matrices.

(a) Show that if $n = 2$, then the Frobenius inner product (see Exercise 10) of A and B is

$$\langle A, B \rangle = a_{11}b_{11} + a_{12}b_{12} + a_{21}b_{21} + a_{22}b_{22}.$$

(b) Extend part (a) to the general case. That is, show that for an arbitrary n ,

$$\langle A, B \rangle = \sum_{i=1}^n \sum_{j=1}^n a_{ij}b_{ij}.$$

(c) Compare the Frobenius inner product to the scalar product of two vectors.

(12) Let $\mathcal{B} = \{[1 \ 1 \ 1]^\top, [1 \ -1 \ 0]^\top\}$ and let $W = \text{Span } \mathcal{B}$ in \mathbb{R}^3 .

(a) Show that \mathcal{B} is an orthogonal basis for W , using the dot product as inner product.

(b) Explain why the vector $\mathbf{v} = [0 \ 2 \ 2]^\top$ is not in W .

(c) Find the vector in W that is closest to \mathbf{v} . How close is this vector to \mathbf{v} ?

(13) Let \mathbb{R}^3 be the inner product space with inner product

$$\langle [u_1 \ u_2 \ u_3]^\top, [v_1 \ v_2 \ v_3]^\top \rangle = u_1v_1 + 2u_2v_2 + u_3v_3.$$

Let $\mathcal{B} = \{[1 \ 1 \ 1]^\top, [1 \ -1 \ 1]^\top\}$ and let $W = \text{Span } \mathcal{B}$ in \mathbb{R}^3 .

(a) Show that \mathcal{B} is an orthogonal basis for W , using the given inner product.

(b) Explain why the vector $\mathbf{v} = [0 \ 2 \ 2]^\top$ is not in W .

(c) Find the vector in W that is closest to \mathbf{v} . How close is this vector to \mathbf{v} ?

(14) Let \mathbb{P}_2 be the inner product space with inner product

$$\langle p(t), q(t) \rangle = \int_0^1 p(t)q(t) \, dt.$$

Let $\mathcal{B} = \{1, 1 - 2t\}$ and let $W = \text{Span } \mathcal{B}$ in \mathbb{P}_2 .

(a) Show that \mathcal{B} is an orthogonal basis for W , using the given inner product.

(b) Explain why the polynomial $q(t) = t^2$ is not in W .

(c) Find the vector in W that is closest to $q(t)$. How close is this vector to $q(t)$?

(15) Prove the remaining properties of Theorem 29.2. That is, if $\langle \ , \ \rangle$ is an inner product on a vector space V and \mathbf{u}, \mathbf{v} , and \mathbf{w} are vectors in V and c is any scalar, then

(a) $\langle \mathbf{0}, \mathbf{v} \rangle = \langle \mathbf{v}, \mathbf{0} \rangle = 0$

(b) $\langle \mathbf{u}, c\mathbf{v} \rangle = c\langle \mathbf{u}, \mathbf{v} \rangle$

(c) $\langle \mathbf{v} + \mathbf{w}, \mathbf{u} \rangle = \langle \mathbf{v}, \mathbf{u} \rangle + \langle \mathbf{w}, \mathbf{u} \rangle$

(d) $\langle \mathbf{u} - \mathbf{v}, \mathbf{w} \rangle = \langle \mathbf{w}, \mathbf{u} - \mathbf{v} \rangle = \langle \mathbf{u}, \mathbf{w} \rangle - \langle \mathbf{v}, \mathbf{w} \rangle = \langle \mathbf{w}, \mathbf{u} \rangle - \langle \mathbf{w}, \mathbf{v} \rangle$



- (16) Prove Theorem 29.16. (Hint: Refer to Theorem 27.10.)
- (17) Prove Theorem 29.8. (Hint: Compare to Theorem 28.3.)
- (18) Let V be a vector space with basis $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$. Define $\langle \cdot, \cdot \rangle$ as follows:

$$\langle \mathbf{u}, \mathbf{w} \rangle = \sum_{i=1}^n u_i w_i$$

if $\mathbf{u} = \sum_{i=1}^n u_i \mathbf{v}_i$ and $\mathbf{w} = \sum_{i=1}^n w_i \mathbf{v}_i$ in V . (Since the representation of a vector as a linear combination of basis elements is unique, this mapping is well-defined.) Show that $\langle \cdot, \cdot \rangle$ is an inner product on V and conclude that any finite dimensional vector space can be made into an inner product space.

- (19) Label each of the following statements as True or False. Provide justification for your response.
- (a) **True/False** The only inner product on \mathbb{R}^n is the dot product.
 - (b) **True/False** If W is a subspace of an inner product space and a vector \mathbf{v} is orthogonal to every vector in a basis of W , then \mathbf{v} is in W^\perp .
 - (c) **True/False** If $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is an orthogonal basis for \mathbb{R}^3 , then so is $\{c\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ for any nonzero scalar c .
 - (d) **True/False** An inner product $\langle \mathbf{u}, \mathbf{v} \rangle$ in an inner product space V results in another vector in V .
 - (e) **True/False** An inner product in an inner product space V is a function that maps pairs of vectors in V to the set of non-negative real numbers.
 - (f) **True/False** The vector space of all $n \times n$ matrices can be made into an inner product space.
 - (g) **True/False** Any non-zero multiple of an inner product on space V is also an inner product on V .
 - (h) **True/False** Every set of k non-zero orthogonal vectors in a vector space V of dimension k is a basis for V .
 - (i) **True/False** For any finite-dimensional inner product space V and a subspace W of V , W is a subspace of $(W^\perp)^\perp$.
 - (j) **True/False** If W is a subspace of an inner product space, then $W \cap W^\perp = \{\mathbf{0}\}$.

Project: Fourier Series and Musical Tones

Joseph Fourier first studied trigonometric polynomials to understand the flow of heat in metallic plates and rods. The resulting series, called Fourier series, now have applications in a variety of areas including electrical engineering, vibration analysis, acoustics, optics, signal processing, image processing, geology, quantum mechanics, and many more. For our purposes, we will focus on synthesized music.



Pure musical tones are periodic sine waves. Simple electronic circuits can be designed to generate alternating current. Alternating current is current that is periodic, and hence is described by a combination of $\sin(kx)$ and $\cos(kx)$ for integer values of k . To synthesize an instrument like a violin, we can project the instrument's tones onto trigonometric polynomials – and then we can produce them electronically. As we will see, these projections are least squares approximations onto certain vector spaces. The website <http://www.falstad.com/fourier/> provides a tool for hearing sounds digitally created by certain functions. For example, you can listen to the sound generated by a sawtooth function f of the form

$$f(x) = \begin{cases} x & \text{if } -\pi < x \leq \pi, \\ f(x - 2\pi), & \text{if } \pi < x, \\ f(x + 2\pi), & \text{if } x \leq -\pi. \end{cases}$$

Try out some of the tones on this website (click on the Sound button to hear the tones). You can also alter the tones by clicking on any one of the white dots and moving it up or down. and play with the buttons. We will learn much about what this website does in this project.

Pure tones are periodic and so are modeled by trigonometric functions. In general, trigonometric polynomials can be used to produce good approximations to periodic phenomena. A *trigonometric polynomial* is an object of the form

$$\begin{aligned} c_0 + c_1 \cos(x) + d_1 \sin(x) + c_2 \cos(2x) + d_2 \sin(2x) + \cdots \\ + c_n \cos(nx) + d_n \sin(nx) + \cdots, \end{aligned}$$

where the c_i and d_j are real constants. With judicious choices of these constants, we can approximate periodic and other behavior with trigonometric polynomials. The first step for us will be to understand the relationships between the summands of these trigonometric polynomials in the inner product space $C[-\pi, \pi]^1$ of continuous functions from $[\pi, \pi]$ to \mathbb{R} with the inner product

$$\langle f, g \rangle = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x)g(x) dx. \quad (29.3)$$

Our first order of business is to verify that (29.3) is, in fact, an inner product.

Project Activity 29.1. Let $C[a, b]$ be the set of continuous real-valued functions on the interval $[a, b]$. In Exercise 1. in Section 29 we are asked to show that $C[a, b]$ is a vector space, while Exercise 2 in Section 29 asks us to show that $\langle f, g \rangle = \int_a^b f(x)g(x) dx$ defines an inner product on $C[a, b]$. However, 29.3 is slightly different than this inner product. Show that any positive scalar multiple of an inner product is an inner product, and conclude that (29.3) defines an inner product on $C[-\pi, \pi]$. (We will see why we introduce the factor of $\frac{1}{\pi}$ later.)

Now we return to our inner product space $C[-\pi, \pi]$ with inner product (29.3). Given a function g in $C[-\pi, \pi]$, we approximate g using only a finite number of the terms in a trigonometric polynomial. Let W_n be the subspace of $C[-\pi, \pi]$ spanned by the functions

$$1, \cos(x), \sin(x), \cos(2x), \sin(2x), \dots, \cos(nx), \sin(nx).$$

One thing we need to know is the dimension of W_n .

¹With suitable adjustments, we can work over any interval that is convenient, but for the sake of simplicity in this project, we will restrict ourselves to the interval $[-\pi, \pi]$.



Project Activity 29.2. We start with the initial case of W_1 .

- (a) Show directly that the functions 1, $\cos(x)$, and $\sin(x)$ are orthogonal.
- (b) What is the dimension of W_1 ? Explain.

Now we need to see if what happened in Project Activity 29.2 happens in general. A few tables of integrals and some basic facts from trigonometry can help.

Project Activity 29.3. A table of integrals shows the following for $k \neq m$ (up to a constant):

$$\int \cos(mx) \cos(kx) dx = \frac{1}{2} \left(\frac{\sin((k-m)x)}{k-m} + \frac{\sin((k+m)x)}{k+m} \right) \quad (29.4)$$

$$\int \sin(mx) \sin(kx) dx = \frac{1}{2} \left(\frac{\sin((k-m)x)}{k-m} - \frac{\sin((k+m)x)}{k+m} \right) \quad (29.5)$$

$$\int \cos(mx) \sin(kx) dx = \frac{1}{2} \left(\frac{\cos((m-k)x)}{m-k} - \frac{\sin((m+k)x)}{m+k} \right) \quad (29.6)$$

$$\int \cos(mx) \sin(mx) dx = -\frac{1}{2m} \cos^2(mx) \quad (29.7)$$

- (a) Use (29.4) to show that $\cos(mx)$ and $\cos(kx)$ are orthogonal in $C[-\pi, \pi]$ if $k \neq m$.
- (b) Use (29.5) to show that $\sin(mx)$ and $\sin(kx)$ are orthogonal in $C[-\pi, \pi]$ if $k \neq m$.
- (c) Use (29.6) to show that $\cos(mx)$ and $\sin(kx)$ are orthogonal in $C[-\pi, \pi]$ if $k \neq m$.
- (d) Use (29.7) to show that $\cos(mx)$ and $\sin(mx)$ are orthogonal in $C[-\pi, \pi]$.
- (e) What is $\dim(W_n)$? Explain.

Once we have an orthogonal basis for W_n , we might want to create an orthonormal basis for W_n . Throughout the remainder of this project, unless otherwise specified, you should use a table of integrals or any appropriate technological tool to find integrals for any functions you need.

Project Activity 29.4. Show that the set

$$\mathcal{B}_n = \left\{ \frac{1}{\sqrt{2}}, \cos(x), \cos(2x), \dots, \cos(nx), \sin(x), \sin(2x), \dots, \sin(nx) \right\}$$

is an orthonormal basis for W_n . Use the fact that the norm of a vector \mathbf{v} in an inner product space with inner product $\langle \cdot, \cdot \rangle$ is defined to be $\sqrt{\langle \mathbf{v}, \mathbf{v} \rangle}$. (This is where the factor of $\frac{1}{\pi}$ will be helpful.)

Now we need to recall how to find the best approximation to a vector by a vector in a subspace, and apply that idea to approximate an arbitrary function g with a trigonometric polynomial in W_n . Recall that the best approximation of a function g in $C[-\pi, \pi]$ is the projection of g onto W_n . If we have an orthonormal basis $\{h_0, h_1, h_2, \dots, h_{2n}\}$ of W_n , then the projection of g onto W_n is

$$\text{proj}_{W_n} g = \langle g, h_0 \rangle h_0 + \langle g, h_1 \rangle h_1 + \langle g, h_2 \rangle h_2 + \dots + \langle g, h_{2n} \rangle h_{2n}.$$

With this idea, we can find formulas for the coefficients when we project an arbitrary function onto W_n .



Project Activity 29.5. If g is an arbitrary function in $C[-\pi, \pi]$, we will write the projection of g onto W_n as

$$a_0 \left(\frac{1}{\sqrt{2}} \right) + a_1 \cos(x) + b_1 \sin(x) + a_2 \cos(2x) + b_2 \sin(2x) + \dots \\ + a_n \cos(nx) + b_n \sin(nx).$$

The a_i and b_j are the *Fourier coefficients* for f . The expression $a_n \cos(nx) + b_n \sin(nx)$ is called the *n th harmonic* of g . The first harmonic is called the *fundamental frequency*. The human ear cannot hear tones whose frequencies exceed 20000 Hz, so we only hear finitely many harmonics (the projections onto W_n for some n).

- (a) Show that

$$a_0 = \frac{1}{\sqrt{2}\pi} \int_{-\pi}^{\pi} g(x) dx. \quad (29.8)$$

Explain why $\frac{a_0}{\sqrt{2}}$ gives the average value of g on $[-\pi, \pi]$. You may want to go back and review average value from calculus. This is saying that the best constant approximation of g on $[-\pi, \pi]$ is its average value, which makes sense.

- (b) Show that for $m \geq 1$,

$$a_m = \frac{1}{\pi} \int_{-\pi}^{\pi} g(x) \cos(mx) dx. \quad (29.9)$$

- (c) Show that for $m \geq 1$,

$$b_m = \frac{1}{\pi} \int_{-\pi}^{\pi} g(x) \sin(mx) dx. \quad (29.10)$$

Let us return to the sawtooth function defined earlier and find its Fourier coefficients.

Project Activity 29.6. Let f be defined by $f(x) = x$ on $[-\pi, \pi]$ and repeated periodically afterwards with period 2π . Let p_n be the projection of f onto W_n .

- (a) Evaluate the integrals to find the projection p_1 .
- (b) Use appropriate technology to find the projections p_{10} , p_{20} , and p_{30} for the sawtooth function f . Draw pictures of these approximations against f and explain what you see.
- (c) Now we find formulas for all the Fourier coefficients. Use the fact that $x \cos(mx)$ is an odd function to explain why $a_m = 0$ for each m . Then show that $b_m = (-1)^{m+1} \frac{2}{m}$ for each m .
- (d) Go back to the website <http://www.falstad.com/fourier/> and replay the sawtooth tone. Explain what the white buttons represent.

Project Activity 29.7. This activity is not connected to the idea of musical tones, so can be safely ignored if so desired. We conclude with a derivation of a very fascinating formula that you may have seen for $\sum_{n=1}^{\infty} \frac{1}{n^2}$. To do so, we need to analyze the error in approximating a function g with a function in W_n .



Let p_n be the projection of g onto W_n . Notice that p_n is also in W_{n+1} . It is beyond the scope of this project, but in “nice” situations we have $\|g - p_n\| \rightarrow 0$ as $n \rightarrow \infty$. Now $g - p_n$ is orthogonal to p_n , so the Pythagorean theorem shows that

$$\|g - p_n\|^2 + \|p_n\|^2 = \|g\|^2.$$

Since $\|g - p_n\|^2 \rightarrow 0$ as $n \rightarrow \infty$, we can conclude that

$$\lim_{n \rightarrow \infty} \|p_n\|^2 = \|g\|^2. \quad (29.11)$$

We use these ideas to derive a formula for $\sum_{n=1}^{\infty} \frac{1}{n^2}$.

- (a) Use the fact that \mathcal{B}_n is an orthonormal basis to show that

$$\|p_n\|^2 = a_0^2 + a_1^2 + b_1^2 + \cdots + a_n^2 + b_n^2.$$

Conclude that

$$\|g\|^2 = a_0^2 + a_1^2 + b_1^2 + \cdots + a_n^2 + b_n^2 + \cdots. \quad (29.12)$$

- (b) For the remainder of this activity, let f be the sawtooth function defined by $f(x) = x$ on $[-\pi, \pi]$ and repeated periodically afterwards. We determined the Fourier coefficients a_i and b_j of this function in Project Activity 29.6.

- i. Show that

$$a_0^2 + a_1^2 + b_1^2 + \cdots + a_n^2 + b_n^2 + \cdots = 4 \sum_{n=1}^{\infty} \frac{1}{n^2}.$$

- ii. Calculate $\|f\|^2$ using the inner product and compare to (29.12) to find a surprising formula for $\sum_{n=1}^{\infty} \frac{1}{n^2}$.

Section 30

The Gram-Schmidt Process

Focus Questions

By the end of this section, you should be able to give precise and thorough answers to the questions listed below. You may want to keep these questions in mind to focus your thoughts as you complete the section.

- What is the Gram-Schmidt process and why is it useful?
- What is the QR-factorization of a matrix and why is it useful?

Application: Gaussian Quadrature

Since integration of functions is difficult, approximation techniques for definite integrals are very important. In calculus we are introduced to various methods, e.g., the midpoint, trapezoid, and Simpson's rule, for approximating definite integrals. These methods divide the interval of integration into subintervals and then use values of the integrand at the endpoints to approximate the integral. These are useful methods when approximating integrals from tabulated data, but there are better methods for other types of integrands. If we make judicious choices in the points we use to evaluate the integrand, we can obtain more accurate results with less work. One such method is Gaussian quadrature (which, for example, is widely used in solving problems of radiation heat transfer in direct integration of the equation of transfer of radiation over space), which we explore later in this section. This method utilizes the Gram-Schmidt process to produce orthogonal polynomials.

Introduction

We have seen that orthogonal bases make computations very convenient. So one question we might want to address is how we can create an orthogonal basis from any basis. We have already done this in the case that $W = \text{Span}\{\mathbf{w}_1\}$ is the span of a single vector in \mathbb{R}^n – we can project the vector \mathbf{v} in the direction of \mathbf{w}_1 . Our goal is to generalize this to project a vector \mathbf{v} onto an entire subspace

in a useful way, and then use that projection to create a vector orthogonal to the subspace.

The Gram-Schmidt Process

To make best approximations by orthogonal projections onto subspaces we need to be able to find orthogonal bases for vector spaces. We now address the question of how to do that.

Preview Activity 30.1. Let $W = \text{Span}\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ in \mathbb{R}^4 (using the dot product), where $\mathbf{v}_1 = [1 \ 1 \ 1 \ 1]^T$, $\mathbf{v}_2 = [-1 \ 4 \ 4 \ -1]^T$, and $\mathbf{v}_3 = [2 \ -1 \ 1 \ 0]^T$. Our goal in this preview activity is to begin to understand how we can find an orthogonal set $\mathcal{B} = \{\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3\}$ in \mathbb{R}^4 so that $\text{Span } \mathcal{B} = W$. To begin, we could start by letting $\mathbf{w}_1 = \mathbf{v}_1$.

- (1) Now we want to find a vector in W that is orthogonal to \mathbf{w}_1 . Let $W_1 = \text{Span}\{\mathbf{w}_1\}$. Explain why $\mathbf{w}_2 = \text{proj}_{\perp W_1} \mathbf{v}_2$ is in W and is orthogonal to \mathbf{w}_1 . Then calculate the vector \mathbf{w}_2 .
- (2) Next we need to find a third vector \mathbf{w}_3 that is in W and is orthogonal to both \mathbf{w}_1 and \mathbf{w}_2 . Let $W_2 = \text{Span}\{\mathbf{w}_1, \mathbf{w}_2\}$. Explain why $\mathbf{w}_3 = \text{proj}_{\perp W_2} \mathbf{v}_3$ is in W and is orthogonal to both \mathbf{w}_1 and \mathbf{w}_2 . Then calculate the vector \mathbf{w}_3 .
- (3) Explain why the set $\{\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3\}$ is an orthogonal basis for W .

Preview Activity 30.1 shows the first steps of the Gram-Schmidt process to construct an orthogonal basis from any basis of a subspace in \mathbb{R}^n . Of course, there is nothing special about using the dot product in Preview Activity 30.1 – the same argument works in any inner product space. To understand why the process works in general, let $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_m\}$ be a basis for a subspace W of an inner product space V . Let $\mathbf{w}_1 = \mathbf{v}_1$ and let $W_1 = \text{Span}\{\mathbf{w}_1\}$. Since $\mathbf{w}_1 = \mathbf{v}_1$ we have that $W_1 = \text{Span}\{\mathbf{w}_1\} = \text{Span}\{\mathbf{v}_1\}$. Now consider the subspace

$$W_2 = \text{Span}\{\mathbf{v}_1, \mathbf{v}_2\}$$

of W . The vectors $\mathbf{v}_1 = \mathbf{w}_1$ and \mathbf{v}_2 are possibly not orthogonal, but we know the orthogonal projection of \mathbf{v}_2 onto W_1^\perp is orthogonal to \mathbf{w}_1 . Let

$$\mathbf{w}_2 = \text{proj}_{\perp W_1} \mathbf{v}_2 = \mathbf{v}_2 - \frac{\langle \mathbf{v}_2, \mathbf{w}_1 \rangle}{\langle \mathbf{w}_1, \mathbf{w}_1 \rangle} \mathbf{w}_1.$$

Then $\{\mathbf{w}_1, \mathbf{w}_2\}$ is an orthogonal set. Note that $\mathbf{w}_1 = \mathbf{v}_1 \neq \mathbf{0}$, and the fact that $\mathbf{v}_2 \notin W_1$ implies that $\mathbf{w}_2 \neq \mathbf{0}$. So the set $\{\mathbf{w}_1, \mathbf{w}_2\}$ is linearly independent, being a set of non-zero orthogonal vectors. Now the question is whether $\text{Span}\{\mathbf{w}_1, \mathbf{w}_2\} = \text{Span}\{\mathbf{v}_1, \mathbf{v}_2\}$. Note that \mathbf{w}_2 is a linear combination of \mathbf{v}_1 and \mathbf{v}_2 , so \mathbf{w}_2 is in $\text{Span}\{\mathbf{v}_1, \mathbf{v}_2\}$. Since $\text{Span}\{\mathbf{w}_1, \mathbf{w}_2\}$ is a 2-dimensional subspace of the 2-dimensional space W_2 , it must be true that $\text{Span}\{\mathbf{w}_1, \mathbf{w}_2\} = W_2 = \text{Span}\{\mathbf{v}_1, \mathbf{v}_2\}$.

Now we take the next step, adding \mathbf{v}_3 into the mix. Let

$$W_3 = \text{Span}\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\} = \text{Span}\{\mathbf{w}_1, \mathbf{w}_2, \mathbf{v}_3\}.$$

The vector

$$\mathbf{w}_3 = \text{proj}_{\perp W_2} \mathbf{v}_3 = \mathbf{v}_3 - \frac{\langle \mathbf{v}_3, \mathbf{w}_1 \rangle}{\langle \mathbf{w}_1, \mathbf{w}_1 \rangle} \mathbf{w}_1 - \frac{\langle \mathbf{v}_3, \mathbf{w}_2 \rangle}{\langle \mathbf{w}_2, \mathbf{w}_2 \rangle} \mathbf{w}_2$$



is orthogonal to both \mathbf{w}_1 and \mathbf{w}_2 and, by construction, \mathbf{w}_3 is a linear combination of \mathbf{v}_1 , \mathbf{v}_2 , and \mathbf{v}_3 . So \mathbf{w}_3 is in W_3 . The fact that $\mathbf{v}_3 \notin W_2$ implies that $\mathbf{w}_3 \neq \mathbf{0}$ and $\{\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3\}$ is a linearly independent set. Since $\text{Span}\{\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3\}$ is a 3-dimensional subspace of the 3-dimensional space W_3 , we conclude that $\text{Span}\{\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3\}$ equals $\text{Span}\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$.

We continue inductively in this same manner. If we have constructed a set $\{\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3, \dots, \mathbf{w}_{k-1}\}$ of $k-1$ orthogonal vectors such that

$$\text{Span}\{\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3, \dots, \mathbf{w}_{k-1}\} = \text{Span}\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \dots, \mathbf{v}_{k-1}\},$$

then we let

$$\begin{aligned}\mathbf{w}_k &= \text{proj}_{\perp W_{k-1}} \mathbf{v}_k \\ &= \mathbf{v}_k - \frac{\langle \mathbf{v}_k, \mathbf{w}_1 \rangle}{\langle \mathbf{w}_1, \mathbf{w}_1 \rangle} \mathbf{w}_1 - \frac{\langle \mathbf{v}_k, \mathbf{w}_2 \rangle}{\langle \mathbf{w}_2, \mathbf{w}_2 \rangle} \mathbf{w}_2 - \cdots - \frac{\langle \mathbf{v}_k, \mathbf{w}_{k-1} \rangle}{\langle \mathbf{w}_{k-1}, \mathbf{w}_{k-1} \rangle} \mathbf{w}_{k-1},\end{aligned}$$

where

$$W_{k-1} = \text{Span}\{\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3, \dots, \mathbf{w}_{k-1}\}.$$

We know that \mathbf{w}_k is orthogonal to $\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_{k-1}$. Since $\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_{k-1}$, and \mathbf{v}_k are all in $W_k = \text{Span}\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k\}$ we see that \mathbf{w}_k is also in W_k . Since $\mathbf{v}_k \notin W_{k-1}$ implies that $\mathbf{w}_k \neq \mathbf{0}$ and $\{\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_k\}$ is a linearly independent set. Then $\text{Span}\{\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3, \dots, \mathbf{w}_k\}$ is a k -dimensional subspace of the k -dimensional space W_k , so it follows that

$$\text{Span}\{\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3, \dots, \mathbf{w}_k\} = W_k = \text{Span}\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \dots, \mathbf{v}_k\}.$$

This process will end when we have an orthogonal set $\{\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3, \dots, \mathbf{w}_m\}$ with $\text{Span}\{\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3, \dots, \mathbf{w}_m\} = W$.

We summarize the process in the following theorem.

Theorem 30.1 (The Gram-Schmidt Process). *Let $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_m\}$ be a basis for a subspace W of an inner product space V . The set $\{\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3, \dots, \mathbf{w}_m\}$ defined by*

- $\mathbf{w}_1 = \mathbf{v}_1$,
- $\mathbf{w}_2 = \mathbf{v}_2 - \frac{\langle \mathbf{v}_2, \mathbf{w}_1 \rangle}{\langle \mathbf{w}_1, \mathbf{w}_1 \rangle} \mathbf{w}_1$,
- $\mathbf{w}_3 = \mathbf{v}_3 - \frac{\langle \mathbf{v}_3, \mathbf{w}_1 \rangle}{\langle \mathbf{w}_1, \mathbf{w}_1 \rangle} \mathbf{w}_1 - \frac{\langle \mathbf{v}_3, \mathbf{w}_2 \rangle}{\langle \mathbf{w}_2, \mathbf{w}_2 \rangle} \mathbf{w}_2$,
- \vdots
- $\mathbf{w}_m = \mathbf{v}_m - \frac{\langle \mathbf{v}_m, \mathbf{w}_1 \rangle}{\langle \mathbf{w}_1, \mathbf{w}_1 \rangle} \mathbf{w}_1 - \frac{\langle \mathbf{v}_m, \mathbf{w}_2 \rangle}{\langle \mathbf{w}_2, \mathbf{w}_2 \rangle} \mathbf{w}_2 - \cdots - \frac{\langle \mathbf{v}_m, \mathbf{w}_{m-1} \rangle}{\langle \mathbf{w}_{m-1}, \mathbf{w}_{m-1} \rangle} \mathbf{w}_{m-1}$.

is an orthogonal basis for W . Moreover,

$$\text{Span}\{\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3, \dots, \mathbf{w}_k\} = \text{Span}\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \dots, \mathbf{v}_k\}$$

for $1 \leq k \leq m$.



The Gram-Schmidt process builds an orthogonal basis $\{\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3, \dots, \mathbf{w}_m\}$ for us from a given basis. To make an orthonormal basis $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3, \dots, \mathbf{u}_m\}$, all we need do is normalize each basis vector: that is, for each i , we let

$$\mathbf{u}_i = \frac{\mathbf{w}_i}{\|\mathbf{w}_i\|}.$$

Activity 30.1. Use the Gram-Schmidt process to find an orthogonal basis for W .

(a) $W = \text{Span} \left\{ \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \right\}$ using the dot product as the inner product

(b) $W = \text{Span}\{1, t, t^2\}$ in \mathbb{P}_2 using the inner product

$$\langle p(t), q(t) \rangle = \int_0^1 p(t)q(t) dt$$

The QR Factorization of a Matrix

There are several different factorizations, or decompositions, of matrices where each matrix is written as a product of certain types of matrices: LU decomposition using lower and upper triangular matrices (see Section 21), EVD (EigenVector Decomposition) decomposition using eigenvectors and diagonal matrices (see Section 18, and in this section we will introduce the QR decomposition using orthogonal matrix and upper triangular matrices. The QR factorization has applications to solving least squares problems and approximating eigenvalues of matrices.

Activity 30.2. Let $A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 2 \end{bmatrix}$.

- (a) Find an orthonormal basis for $\text{Col } A$. Let \mathcal{B} be the basis formed by these vectors. Let Q be the matrix whose columns are these orthonormal basis vectors.
- (b) Write the columns of A as linear combinations of the columns of Q . That is, if $A = [\mathbf{a}_1 \ \mathbf{a}_2]$, find $[\mathbf{a}_1]_{\mathcal{B}}$ and $[\mathbf{a}_2]_{\mathcal{B}}$. Let $R = [[\mathbf{a}_1]_{\mathcal{B}} \ [\mathbf{a}_2]_{\mathcal{B}}]$.
- (c) Find the product QR and compare to A .

Activity 30.2 contains the main ideas to find the QR factorization of a matrix. Let

$$A = [\mathbf{a}_1 \ | \ \mathbf{a}_2 \ | \ \mathbf{a}_3 \ | \ \cdots \ | \ \mathbf{a}_n]$$

be an $m \times n$ matrix with rank¹ n . We can use the Gram-Schmidt process to find an orthonormal basis $\{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n\}$ for $\text{Col } A$. Recall also that $\text{Span}\{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k\} = \text{Span}\{\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_k\}$ for any k between 1 and n . Let

$$Q = [\mathbf{u}_1 \ \mathbf{u}_2 \ \mathbf{u}_3 \ \cdots \ \mathbf{u}_n].$$

If k is between 1 and n , then \mathbf{a}_k is in $\text{Span}\{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k\}$ and

$$a_k = r_{k1}\mathbf{u}_1 + r_{k2}\mathbf{u}_2 + \cdots + r_{kk}\mathbf{u}_k$$

¹Recall that the rank of a matrix A is the dimension of the column space of A .



for some scalars $r_{k1}, r_{k2}, \dots, r_{kk}$. Then

$$Q \begin{bmatrix} r_{k1} \\ r_{k2} \\ \vdots \\ r_{kk} \\ 0 \\ \vdots \\ 0 \end{bmatrix} = \mathbf{a}_k.$$

If we let $\mathbf{r}_k = \begin{bmatrix} r_{k1} \\ r_{k2} \\ \vdots \\ r_{kk} \\ 0 \\ \vdots \\ 0 \end{bmatrix}$ for k from 1 to n , then

$$A = [Q\mathbf{r}_1 \ Q\mathbf{r}_2 \ Q\mathbf{r}_3 \ \cdots \ Q\mathbf{r}_n].$$

This is the QR factorization of A into the product

$$A = QR$$

where the columns of Q form an orthonormal basis for $\text{Col } A$ and

$$R = [\mathbf{r}_1 \ \mathbf{r}_2 \ \mathbf{r}_3 \ \cdots \ \mathbf{r}_n] = \begin{bmatrix} r_{11} & r_{12} & r_{13} & \cdots & r_{1n-1} & r_{1n} \\ 0 & r_{22} & r_{23} & \cdots & r_{2n-1} & r_{2n} \\ 0 & 0 & r_{33} & \cdots & r_{3n-1} & r_{3n} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & r_{nn} \end{bmatrix}$$

is an upper triangular matrix. Note that Q is an $m \times n$ matrix and R is an $n \times n$ matrix.

Activity 30.3. Let $A = \begin{bmatrix} 1 & 0 & 2 \\ 0 & 2 & 0 \end{bmatrix}$. Find the QR factorization of the following matrices or explain why the matrix does not have a QR factorization.

The QR factorization provides a widely used algorithm (the QR algorithm) for approximating all of the eigenvalues of a matrix. The computer system MATLAB utilizes four versions of the QR algorithm to approximate the eigenvalues of real symmetric matrices, eigenvalues of real nonsymmetric matrices, eigenvalues of pairs of complex matrices, and singular values of general matrices.

The algorithm works as follows.

- Start with an $n \times n$ matrix A . Let $A_1 = A$.



- Find the QR factorization for A_1 and write it as $A_1 = Q_1 R_1$, where Q_1 is orthogonal and R_1 is upper triangular.
- Let $A_2 = Q_1^{-1} A_1 Q_1 = Q_1^T A Q_1 = R_1 Q_1$. Find the QR factorization of A_2 and write it as $A_2 = Q_2 R_2$.
- Let $A_3 = Q_2^{-1} A_2 Q_2 = Q_2^T A Q_2 = R_2 Q_2$. Find the QR factorization of A_3 and write it as $A_3 = Q_3 R_3$.
- Continue in this manner to obtain a sequence $\{A_k\}$ where $A_k = Q_k R_k$ and $A_{k+1} = R_k Q_k$.

Note that $A_{k+1} = Q_k^{-1} A_k Q_k$ and so all of the matrices A_k are similar to each other and therefore all have the same eigenvalues. We won't go into the details, but it can be shown that if the eigenvalues of A are real and have distinct absolute values, then the sequence $\{A_i\}$ converges to an upper triangular matrix with the eigenvalues of A as the diagonal entries. If some of the eigenvalues of A are complex, then the sequence $\{A_i\}$ converges to a block upper triangular matrix, where the diagonal blocks are either 1×1 (approximating the real eigenvalues of A) or 2×2 (which provide a pair of conjugate complex eigenvalues of A).

Examples

What follows are worked examples that use the concepts from this section.

Example 30.2. Let W be the subspace of \mathbb{R}^4 spanned by $\mathbf{w}_1 = [1 \ 0 \ 0 \ 0]^T$, $\mathbf{w}_2 = [1 \ 1 \ 1 \ 0]^T$, and $\mathbf{w}_3 = [1 \ 2 \ 0 \ 1]^T$.

- (a) Use the Gram-Schmidt process to find an orthonormal basis for the subspace W of \mathbb{R}^4 spanned by $\mathbf{w}_1 = [1 \ 0 \ 0 \ 0]^T$, $\mathbf{w}_2 = [1 \ 1 \ 1 \ 0]^T$, and $\mathbf{w}_3 = [1 \ 2 \ 0 \ 1]^T$.

(b) Find a QR factorization of the matrix $A = \begin{bmatrix} 1 & 1 & 1 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix}$.

Example Solution.

- (a) First note that \mathbf{w}_1 , \mathbf{w}_2 , and \mathbf{w}_3 are linearly independent. We let $\mathbf{v}_1 = \mathbf{w}_1$ and the Gram-Schmidt process gives us

$$\begin{aligned} \mathbf{v}_2 &= \mathbf{w}_2 - \frac{\mathbf{w}_2 \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1} \mathbf{v}_1 \\ &= [1 \ 1 \ 1 \ 0]^T - \frac{1}{1}[1 \ 0 \ 0 \ 0]^T \\ &= [0 \ 1 \ 1 \ 0]^T \end{aligned}$$

and

$$\begin{aligned} \mathbf{v}_3 &= \mathbf{w}_3 - \frac{\mathbf{w}_3 \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1} \mathbf{v}_1 - \frac{\mathbf{w}_3 \cdot \mathbf{v}_2}{\mathbf{v}_2 \cdot \mathbf{v}_2} \mathbf{v}_2 \\ &= [1 \ 2 \ 0 \ 1]^T - \frac{1}{1}[1 \ 0 \ 0 \ 0]^T - \frac{2}{2}[0 \ 1 \ 1 \ 0]^T \\ &= [0 \ 1 \ -1 \ 1]^T. \end{aligned}$$

So $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is an orthogonal basis for W . An orthonormal basis is found by dividing each vector by its magnitude, so

$$\{[1 0 0 0]^\top, \frac{1}{\sqrt{2}}[0 1 1 0]^\top, \frac{1}{\sqrt{3}}[0 1 -1 1]^\top\}$$

is an orthonormal basis for W .

- (b) Technology shows that the reduced row echelon form of A is I_4 , so the columns of A are linearly independent and A has rank 4. From part (a) we have an orthogonal basis for the span of the first three columns of A . To find a fourth vector to add so that the span is $\text{Col } A$, we apply the Gram-Schmidt process one more time with $\mathbf{w}_4 = [0 0 0 1]^\top$:

$$\begin{aligned}\mathbf{v}_4 &= \mathbf{w}_4 - \frac{\mathbf{w}_4 \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1} \mathbf{v}_1 - \frac{\mathbf{w}_4 \cdot \mathbf{v}_2}{\mathbf{v}_2 \cdot \mathbf{v}_2} \mathbf{v}_2 - \frac{\mathbf{w}_4 \cdot \mathbf{v}_3}{\mathbf{v}_3 \cdot \mathbf{v}_3} \mathbf{v}_3 \\ &= [0 0 0 1]^\top - \frac{0}{1}[1 0 0 0]^\top - \frac{0}{2}[0 1 1 0]^\top - \frac{1}{3}[0 1 -1 1]^\top \\ &= \frac{1}{3}[0 -1 1 2]^\top.\end{aligned}$$

So $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3, \mathbf{u}_4\}$ is an orthonormal basis for $\text{Col } A$, where

$$\begin{aligned}\mathbf{u}_1 &= [1 0 0 0]^\top & \mathbf{u}_2 &= \frac{\sqrt{2}}{2}[0 1 1 0]^\top \\ \mathbf{u}_3 &= \frac{\sqrt{3}}{3}[0 1 -1 1]^\top & \mathbf{u}_4 &= \frac{\sqrt{6}}{6}[0 -1 1 2]^\top.\end{aligned}$$

This makes

$$Q = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \frac{\sqrt{2}}{2} & \frac{\sqrt{3}}{3} & -\frac{\sqrt{6}}{6} \\ 0 & \frac{\sqrt{2}}{2} & -\frac{\sqrt{3}}{3} & \frac{\sqrt{6}}{6} \\ 0 & 0 & \frac{\sqrt{3}}{3} & \frac{\sqrt{6}}{3} \end{bmatrix}.$$

To find the matrix R , we write the columns of A in terms of the basis vectors $\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3$, and \mathbf{u}_4 . Technology shows that the reduced row echelon form of $[Q \mid A]$ is

$$\left[\begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & 1 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & \sqrt{2} & \sqrt{2} & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & \sqrt{3} & \frac{\sqrt{3}}{3} \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & \frac{\sqrt{6}}{3} \end{array} \right].$$

So

$$R = \begin{bmatrix} 1 & 1 & 1 & 0 \\ 0 & \sqrt{2} & \sqrt{2} & 0 \\ 0 & 0 & \sqrt{3} & \frac{\sqrt{3}}{3} \\ 0 & 0 & 0 & \frac{\sqrt{6}}{3} \end{bmatrix}.$$

Example 30.3. Consider the space $V = C[0, 1]$ with inner product $\langle f, g \rangle = \int_0^1 f(x)g(x) dx$.



- (a) Find the polynomial in \mathbb{P}_2 (considered as a subspace of V) that is closest to $h(x) = \sqrt{x}$. Use technology to calculate any required integrals. Draw a graph of your approximation against the graph of h .
- (b) Provide a numeric measure of the error in approximating \sqrt{x} by the polynomial you found in part (a).

Example Solution.

- (a) Our job is to find $\text{proj}_{\mathbb{P}_2} h(x)$. To do this, we need an orthogonal basis of \mathbb{P}_2 . We apply the Gram-Schmidt process to the standard basis $\{1, t, t^2\}$ of \mathbb{P}_2 to obtain an orthogonal basis $\{p_1(t), p_2(t), p_3(t)\}$ of \mathbb{P}_2 . We start with $p_1(t) = 1$, then

$$\begin{aligned} p_2(t) &= t - \frac{\langle t, 1 \rangle}{\langle 1, 1 \rangle}(1) \\ &= t - \frac{1}{2} \end{aligned}$$

and

$$\begin{aligned} p_3(t) &= t^2 - \frac{\langle t^2, 1 \rangle}{\langle 1, 1 \rangle}(1) - \frac{\langle t^2, t - \frac{1}{2} \rangle}{\langle t - \frac{1}{2}, t - \frac{1}{2} \rangle} \left(t - \frac{1}{2} \right) \\ &= t^2 - \frac{1}{3} - \frac{\frac{1}{12}}{\frac{1}{12}} \left(t - \frac{1}{2} \right) \\ &= t^2 - t + \frac{1}{6}. \end{aligned}$$

Then

$$\begin{aligned} \text{proj}_{\mathbb{P}_2} \sqrt{x} &= \frac{\langle \sqrt{x}, 1 \rangle}{\langle 1, 1 \rangle}(1) + \frac{\langle \sqrt{x}, x - \frac{1}{2} \rangle}{\langle x - \frac{1}{2}, x - \frac{1}{2} \rangle} \left(x - \frac{1}{2} \right) \\ &\quad + \frac{\langle \sqrt{x}, x^2 - x + \frac{1}{12} \rangle}{\langle x^2 - x + \frac{1}{6}, x^2 - x + \frac{1}{6} \rangle} \left(x^2 - x + \frac{1}{6} \right) \\ &= \frac{\frac{2}{3}}{1}(1) + \frac{\frac{1}{15}}{\frac{1}{12}} \left(x - \frac{1}{2} \right) - \frac{\frac{1}{315}}{\frac{1}{180}} \left(x^2 - x + \frac{1}{6} \right) \\ &= \frac{2}{3} + \left(\frac{4}{5} \right) \left(x - \frac{1}{2} \right) - \frac{4}{7} \left(x^2 - x + \frac{1}{6} \right) \\ &= -\frac{4}{7}x^2 + \frac{48}{35}x + \frac{6}{35}. \end{aligned}$$

A graph of the approximation and h are shown in Figure 30.1

- (b) The norm of $\text{proj}_{\mathbb{P}_2^\perp} \sqrt{x} = \sqrt{x} - \left(-\frac{4}{7}x^2 + \frac{48}{35}x + \frac{6}{35} \right)$ tells us how well our projection $-\frac{4}{7}x^2 + \frac{48}{35}x + \frac{6}{35}$ approximates \sqrt{x} . Technology shows that

$$\|\text{proj}_{\mathbb{P}_2^\perp} \sqrt{x}\| = \sqrt{\langle \text{proj}_{\mathbb{P}_2^\perp} \sqrt{x}, \text{proj}_{\mathbb{P}_2^\perp} \sqrt{x} \rangle} = \frac{\sqrt{2}}{70} \approx 0.02.$$



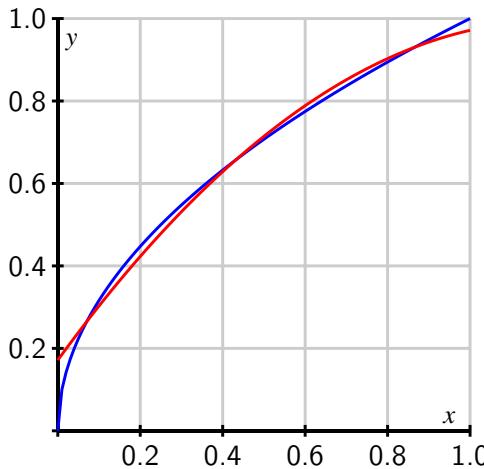


Figure 30.1: The graphs of \sqrt{x} and $-\frac{4}{7}x^2 + \frac{48}{35}x + \frac{6}{35}$.

Summary

- The Gram-Schmidt process produces an orthogonal basis from any basis. It works as follows. Let $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_m\}$ be a basis for a subspace W of an inner product space V . The set $\{\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3, \dots, \mathbf{w}_m\}$ defined by

$$\begin{aligned} & - \mathbf{w}_1 = \mathbf{v}_1, \\ & - \mathbf{w}_2 = \mathbf{v}_2 - \frac{\langle \mathbf{v}_2, \mathbf{w}_1 \rangle}{\langle \mathbf{w}_1, \mathbf{w}_1 \rangle} \mathbf{w}_1, \\ & - \mathbf{w}_3 = \mathbf{v}_3 - \frac{\langle \mathbf{v}_3, \mathbf{w}_1 \rangle}{\langle \mathbf{w}_1, \mathbf{w}_1 \rangle} \mathbf{w}_1 - \frac{\langle \mathbf{v}_3, \mathbf{w}_2 \rangle}{\langle \mathbf{w}_2, \mathbf{w}_2 \rangle} \mathbf{w}_2, \\ & \vdots \\ & - \mathbf{w}_m = \mathbf{v}_m - \frac{\langle \mathbf{v}_m, \mathbf{w}_1 \rangle}{\langle \mathbf{w}_1, \mathbf{w}_1 \rangle} \mathbf{w}_1 - \frac{\langle \mathbf{v}_m, \mathbf{w}_2 \rangle}{\langle \mathbf{w}_2, \mathbf{w}_2 \rangle} \mathbf{w}_2 - \cdots - \frac{\langle \mathbf{v}_m, \mathbf{w}_{m-1} \rangle}{\langle \mathbf{w}_{m-1}, \mathbf{w}_{m-1} \rangle} \mathbf{w}_{m-1}. \end{aligned}$$

is an orthogonal basis for W . Moreover,

$$\text{Span}\{\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3, \dots, \mathbf{w}_k\} = \text{Span}\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \dots, \mathbf{v}_k\}$$

for each k between 1 and m .

- The QR factorization has applications to solving least squares problems and approximating eigenvalues of matrices. The QR factorization writes an $m \times n$ matrix with rank n as a product $A = QR$, where the columns of Q form an orthonormal basis for $\text{Col } A$ and

$$R = [\mathbf{r}_1 \mid \mathbf{r}_2 \mid \mathbf{r}_3 \mid \cdots \mid \mathbf{r}_n] = \begin{bmatrix} r_{11} & r_{12} & r_{13} & \cdots & r_{1n-1} & r_{1n} \\ 0 & r_{22} & r_{23} & \cdots & r_{2n-1} & r_{2n} \\ 0 & 0 & r_{33} & \cdots & r_{3n-1} & r_{3n} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & r_{nn} \end{bmatrix}$$

is an upper triangular matrix.

Exercises

(1) Let $\mathbf{w}_1 = [1 \ 2 \ 1]^\top$ and $\mathbf{w}_2 = [1 \ -1 \ 1]^\top$, and Let $W = \text{Span}\{\mathbf{w}_1, \mathbf{w}_2\}$ in \mathbb{R}^3 with the dot product as inner product. Let $\mathbf{v} = [1 \ 0 \ 0]^\top$.

(a) Find $\text{proj}_W \mathbf{v}$ and $\text{proj}_{W^\perp} \mathbf{v}$

(b) Find the vector in W that is closest to \mathbf{v} . How close is this vector to \mathbf{v} ?

(2) Let $\mathcal{B} = \left\{ \begin{bmatrix} 1 \\ 0 \\ -1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \\ -1 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \\ -1 \\ 1 \end{bmatrix} \right\}$ and let $W = \text{Span}(\mathcal{B})$ in \mathbb{R}^4 . Find the best

approximation in W to the vector $\mathbf{v} = \begin{bmatrix} 2 \\ 3 \\ 1 \\ -1 \end{bmatrix}$ in W and give a numeric estimate of how good this approximation is.

(3) Let $V = C[0, 1]$ with the inner product $\langle f(t), g(t) \rangle = \int_0^1 f(t)g(t) dt$. Let $W = \mathbb{P}_2$. Note that W is a subspace of V .

(a) Find the polynomial $q(t)$ in W that is closest to the function h defined by $h(t) = \frac{2}{1+t^2}$ in the least squares sense. That is, find the projection of $h(t)$ onto W . (Hint: Recall the work done in Activity 30.1.)

(b) Find the second order Taylor polynomial $P_2(t)$ for $h(t)$ centered at 0.

(c) Plot $h(t)$, $q(t)$, and $P_2(t)$ on the same set of axes. Which approximation provides a better fit to h on this interval. Why?

(4) In this exercise we determine the least-squares line (the line of best fit in the least squares sense) to a set of k data points $(x_1, y_1), (x_2, y_2), \dots, (x_k, y_k)$ in the plane. In this case, we want to fit a line of the form $f(x) = mx + b$ to the data. If the data were to lie on a line, then we would have a solution to the system

$$mx_1 + b = y_1$$

$$mx_2 + b = y_2$$

\vdots

$$mx_k + b = y_k$$

This system can be written as

$$m\mathbf{w}_1 + b\mathbf{w}_2 = \mathbf{y},$$

where $\mathbf{w}_1 = [x_1 \ x_2 \ \cdots \ x_k]^\top$, $\mathbf{w}_2 = [1 \ 1 \ \cdots \ 1]^\top$, and $\mathbf{y} = [y_1 \ y_2 \ \cdots \ y_k]^\top$. If the data does not lie on a line, then the system won't have a solution. Instead, we want to minimize the square of the distance between \mathbf{y} and a vector of the form $m\mathbf{w}_1 + b\mathbf{w}_2$. That is, minimize

$$\|\mathbf{y} - (m\mathbf{w}_1 + b\mathbf{w}_2)\|^2. \quad (30.1)$$



Rephrasing this in terms of projections, we are looking for the vector in $W = \text{Span}\{\mathbf{w}_1, \mathbf{w}_2\}$ that is closest to \mathbf{y} . In other words, the values of m and b will occur as the weights when we write $\text{proj}_W \mathbf{y}$ as a linear combination of \mathbf{w}_1 and \mathbf{w}_2 . The one wrinkle in this problem is that we need an orthogonal basis for W to find this projection. Find the least squares line for the data points $(1, 2)$, $(2, 4)$ and $(3, 5)$ in \mathbb{R}^2 .

- (5) Each set S is linearly independent. Use the Gram-Schmidt process to create an orthogonal set of vectors with the same span as S . Then find an orthonormal basis for the same span.

(a) $S = \left\{ \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 5 \\ -1 \\ 2 \end{bmatrix} \right\}$ in \mathbb{R}^3 using the dot product as the inner product

(b) $S = \{1 + t, 1 - t, t - t^2\}$ in \mathbb{P}_2 using the inner product

$$\langle p(t), q(t) \rangle = \int_{-1}^1 p(t)q(t) dt$$

(c) $S = \left\{ \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} \right\}$ in \mathbb{R}^3 using the inner product $\langle \mathbf{u}, \mathbf{v} \rangle = (\mathbf{A}\mathbf{u}) \cdot (\mathbf{A}\mathbf{v})$, where $A = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 0 \end{bmatrix}$

(d) $S = \left\{ \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \\ 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 2 \\ 3 \\ 0 \\ 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 4 \\ -1 \\ 2 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \right\}$ in \mathbb{R}^7 with the weighted inner product

$$\begin{aligned} \langle [u_1 u_2 u_3 u_4 u_5 u_6 u_7]^T, [v_1 v_2 v_3 v_4 v_5 v_6 v_7]^T \rangle \\ = u_1 v_1 + u_2 v_2 + u_3 v_3 + 2u_4 v_4 + 2u_5 v_5 + u_6 v_6 + u_7 v_7 \end{aligned}$$

- (6) Let $S = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ be a set of linearly dependent vectors in an inner product space V . What is the result if the Gram-Schmidt process is applied to the set S ? Explain.

- (7) Let $S = \{1, \cos(t), \sin(t)\}$.

- (a) Show that S is a linearly independent set in $C[0, \pi]$.

- (b) Use the Gram-Schmidt process to find an orthogonal basis from S using the inner product

$$\langle f(t), g(t) \rangle = \int_0^\pi f(t)g(t) dt$$

- (8) The Legendre polynomials form an orthonormal basis for the infinite dimensional inner product space \mathbb{P} of all polynomials using the inner product

$$\langle p(t), q(t) \rangle = \int_{-1}^1 p(t)q(t) dt.$$



The Legendre polynomials have applications to differential equations, statistics, numerical analysis, and physics (e.g., they appear when solving Schrödinger equation in three dimensions for a central force). The Legendre polynomials are found by using the Gram-Schmidt process to find an orthogonal basis from the standard basis $\{1, t, t^2, \dots\}$ for \mathbb{P} . Find the first four Legendre polynomials by creating a orthonormal basis from the set $\{1, t, t^2, t^3\}$.

- (9) A fellow student wants to find a QR factorization for a 3×4 matrix. What would you tell this student and why?
- (10) Label each of the following statements as True or False. Provide justification for your response. Throughout, let V be a vector space.
 - (a) **True/False** If $\{\mathbf{w}_1, \mathbf{w}_2\}$ is a basis for a subspace W of an inner product space V , then the vector $\frac{\mathbf{v} \cdot \mathbf{w}_1}{\mathbf{w}_1 \cdot \mathbf{w}_1} \mathbf{w}_1 + \frac{\mathbf{v} \cdot \mathbf{w}_2}{\mathbf{w}_2 \cdot \mathbf{w}_2} \mathbf{w}_2$ is the vector in W closest to \mathbf{v} .
 - (b) **True/False** If W is a subspace of an inner product space V , then the vector $\mathbf{v} - \text{proj}_W \mathbf{v}$ is orthogonal to every vector in W .
 - (c) **True/False** If $\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3$ are vectors in an inner product space V , then the Gram-Schmidt process constructs an orthogonal set of vectors $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ with the same span as $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$.
 - (d) **True/False** Any set $\{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k\}$ of orthogonal vectors in an inner product space V can be extended to an orthogonal basis of V .
 - (e) **True/False** If A is an $n \times n$ matrix with $AA^\top = I_n$, then the rows of A form an orthogonal set.
 - (f) **True/False** Every nontrivial finite dimensional subspace of an inner product space has an orthogonal basis.
 - (g) **True/False** In any inner product space V , if W is a subspace satisfying $W^\perp = \{\mathbf{0}\}$, then $W = V$.

Project: Gaussian Quadrature and Legendre Polynomials

Simpson's rule is a reasonably accurate method for approximating definite integrals since it models the integrand on subintervals with quadratics. For that reason, Simpson's rule provides exact values for integrals of all polynomials of degree less than or equal to 2. In Gaussian quadrature, we will use a family of polynomials to determine points at which to evaluate an integral of the form $\int_{-1}^1 f(t) dt$. By allowing ourselves to select evaluation points that are not uniformly distributed across the interval of integration, we will be able to approximate our integrals much more efficiently. The method is constructed so as to obtain exact values for as large of degree polynomial integrands as possible. As a result, if we can approximate our integrand well with polynomials, we can obtain very good approximations with Gaussian quadrature with minimal effort.

The Gaussian quadrature approximation has the form

$$\int_{-1}^1 f(t) dt \approx w_1 f(t_1) + w_2 f(t_2) + \dots + w_n f(t_n) = \sum_{i=1}^n w_i f(t_i), \quad (30.2)$$



where the w_i (weights) and the t_i (nodes) are points in the interval $[-1, 1]^2$. Gaussian quadrature describes how to find the weights and the points in (30.2) to obtain suitable approximations. We begin to explore Gaussian quadrature with the simplest cases.

Project Activity 30.1. In this activity we find through direct calculation the node and weight with $n = 1$ so that

$$w_1 f(t_1) \approx \int_{-1}^1 f(t) dt. \quad (30.3)$$

There are two unknowns in this situation (w_1 and t_1) and so we will need 2 equations to find these unknowns. Keep in mind that we want to have the approximation (30.3) be exact for as large of degree polynomials as possible.

- (a) Assume equality in (30.3) if we choose $f(t) = 1$. Use the resulting equation to find w_1 .
- (b) Assume equality in (30.3) if we choose $f(t) = t$. Use the resulting equation to find t_1 .
- (c) Verify that (30.3) is in fact an equality for any linear polynomial of the form $f(t) = a_0 + a_1 t$, using the values of w_1 and t_1 you found

We do one more specific case before considering the general case.

Project Activity 30.2. In this problem we find through direct calculation the nodes and weights with $n = 2$ so that

$$w_1 f(t_1) + w_2 f(t_2) \approx \int_{-1}^1 f(t) dt. \quad (30.4)$$

There are four unknowns in this situation (w_1, w_2 and t_1, t_2) and so we will need 4 equations to find these unknowns. Keep in mind that we want to have the approximation (30.4) be exact for as large of degree polynomials as possible. In this case we will use $f(t) = 1$, $f(t) = t$, $f(t) = t^2$, and $f(t) = t^3$.

- (a) Assume equality in (30.4) if we choose $f(t) = 1$. This gives us an equation in w_1 and w_2 . Find this equation.
- (b) Assume equality in (30.4) if we choose $f(t) = t$. This gives us an equation in w_1, w_2 and t_1, t_2 . Find this equation.
- (c) Assume equality in (30.4) if we choose $f(t) = t^2$. This gives us an equation in w_1, w_2 and t_1, t_2 . Find this equation.
- (d) Assume equality in (30.4) if we choose $f(t) = t^3$. This gives us an equation in w_1, w_2 and t_1, t_2 . Find this equation.
- (e) Solve this system of 4 equations in 4 unknowns. You can do this by hand or with any other appropriate tool. Show that t_1 and t_2 are the roots of the polynomial $t^2 - \frac{1}{3}$.
- (f) Verify that (30.4) is in fact an equality with the values of w_1, w_2 and t_1, t_2 you found for any polynomial of the form $f(t) = a_0 + a_1 t + a_2 t^2 + a_3 t^3$.

²As we will see later, integrations over $[a, b]$ can be converted to an integral over $[-1, 1]$ with a change of variable.



Other than solving a system of linear equations as in Project Activity 30.2, it might be reasonable to ask what the connection is between Gaussian quadrature and linear algebra. We explore that connection now.

In the general case, we want to find the weights and nodes to make the approximation exact for as large degree polynomials as possible. We have $2n$ unknowns w_1, w_2, \dots, w_n and t_1, t_2, \dots, t_n , so we need to impose $2n$ conditions to determine the unknowns. We will require equality for the $2n$ functions t^i for i from 0 to $2n - 1$. This yields the equations

$$\begin{aligned} w_1 \cdot 1 + w_2 \cdot 1 + \cdots + w_n \cdot 1 &= \int_{-1}^1 1 \, dt = t \Big|_{-1}^1 = 2 \\ w_1 t_1 + w_2 t_2 + \cdots + w_n t_n &= \int_{-1}^1 t \, dt = \frac{t^2}{2} \Big|_{-1}^1 = 0 \\ w_1 t_1^2 + w_2 t_2^2 + \cdots + w_n t_n^2 &= \int_{-1}^1 t^2 \, dt = \frac{t^3}{3} \Big|_{-1}^1 = \frac{2}{3} \\ &\vdots \quad \vdots \\ w_1 t_1^{2n-1} + w_2 t_2^{2n-1} + \cdots + w_n t_n^{2n-1} &= \int_{-1}^1 t^{2n-1} \, dt = \frac{t^{2n}}{2n} \Big|_{-1}^1 = 0. \end{aligned}$$

In the i th equation the right hand side is

$$\int_{-1}^1 t^i \, dt = \frac{t^{i+1}}{i+1} \Big|_{-1}^1 = \begin{cases} \frac{2}{i+1} & \text{if } i \text{ is even,} \\ 0 & \text{if } i \text{ is odd.} \end{cases}$$

Project Activity 30.3. It is inefficient to always solve these systems of equations to find the nodes and weights, especially since there is a more elegant way to find the nodes.

- (a) Use appropriate technology to find the equations satisfied by the t_i for $n = 3$, $n = 4$, and $n = 5$.
- (b) Now we will see the more elegant way to find the nodes. As we will show for some cases, the nodes can be found as roots of a set of orthogonal polynomials in \mathbb{P}_n with the inner product $\langle f(t), g(t) \rangle = \int_{-1}^1 f(t)g(t) \, dt$. Begin with the basis $\mathcal{S}_n = \{1, t, t^2, \dots, t^n\}$ of \mathbb{P}_n . Use appropriate technology to find an orthogonal basis B_5 for \mathbb{P}_5 obtained by applying the Gram-Schmidt process to \mathcal{S}_n . The polynomials in this basis are called *Legendre polynomials*. Check that the nodes are roots of the Legendre polynomials by finding roots of these polynomials using any method. Explain why the t_i appear to be roots of the Legendre polynomials.

Although it would take us beyond the scope of this project to verify this fact, the nodes in the n th Gaussian quadrature approximation (30.2) are in fact the roots of the n th order Legendre polynomial. In other words, if $p_n(t)$ is the n th order Legendre polynomial, then t_1, t_2, \dots, t_n are the roots of $p_n(t)$ in $[-1, 1]$. Gaussian quadrature as described in (30.2) using the polynomial $p_n(t)$ is exact if the integrand $f(t)$ is a polynomial of degree less than $2n$.



We can find the corresponding weights, w_i , using the formula³

$$w_i = \frac{2}{(1 - t_i^2)(q'_n(t_i))^2}, \quad (30.5)$$

where $q_i(t)$ is the i th order Legendre polynomial scaled so that $q_i(1) = 1$.

Project Activity 30.4. Let us see now how good the integral estimates are with Gaussian quadrature method using an example. Use Gaussian quadrature with the indicated value of n to approximate $\int_{-1}^1 e^t \cos(t) dt$. Be sure to explain how you found your nodes and weights (approximate the nodes and weights to 8 decimal places). Compare the approximations with the actual value of the integral. Use technology as appropriate to help with calculations.

- (a) $n = 3$
- (b) $n = 4$
- (c) $n = 5$

Our Gaussian quadrature formula was derived for integrals on the interval $[-1, 1]$. We conclude by seeing how a definite integral on an interval can be converted to one on the interval $[-1, 1]$.

Project Activity 30.5. Consider the problem of approximating an integral of the form $I = \int_a^b g(x) dx$. Show that the change of variables $x = \frac{(b-a)t}{2} + \frac{a+b}{2}$, $f(t) = \frac{(b-a)g(x)}{2}$ reduces the integral I to the form $I = \int_{-1}^1 f(t) dt$. (This change of variables can be derived by finding a linear function that maps the interval $[a, b]$ to the interval $[-1, 1]$.)

³Abramowitz, Milton; Stegun, Irene A., eds. (1972), 25.4, Integration, Handbook of Mathematical Functions (with Formulas, Graphs, and Mathematical Tables), Dover,





Part VII

Orthogonality

Section 31

Orthogonal Diagonalization

Focus Questions

By the end of this section, you should be able to give precise and thorough answers to the questions listed below. You may want to keep these questions in mind to focus your thoughts as you complete the section.

- What does it mean for a matrix to be orthogonally diagonalizable and why is this concept important?
- What is a symmetric matrix and what important property related to diagonalization does a symmetric matrix have?
- What is the spectrum of a matrix?

Application: The Multivariable Second Derivative Test

In single variable calculus, we learn that the second derivative can be used to classify a critical point of the type where the derivative of a function is 0 as a local maximum or minimum.

Theorem 31.1 (The Second Derivative Test for Single-Variable Functions). *If a is a critical number of a function f so that $f'(a) = 0$ and if $f''(a)$ exists, then*

- if $f''(a) < 0$, then $f(a)$ is a local maximum value of f ,
- if $f''(a) > 0$, then $f(a)$ is a local minimum value of f , and
- if $f''(a) = 0$, this test yields no information.

In the two-variable case we have an analogous test, which is usually seen in a multivariable calculus course.

Theorem 31.2 (The Second Derivative Test for Functions of Two Variables). *Suppose (a, b) is a critical point of the function f for which $f_x(a, b) = 0$ and $f_y(a, b) = 0$. Let D be the quantity*

defined by

$$D = f_{xx}(a, b)f_{yy}(a, b) - f_{xy}(a, b)^2.$$

- If $D > 0$ and $f_{xx}(a, b) < 0$, then f has a local maximum at (a, b) .
- If $D > 0$ and $f_{xx}(a, b) > 0$, then f has a local minimum at (a, b) .
- If $D < 0$, then f has a saddle point at (a, b) .
- If $D = 0$, then this test yields no information about what happens at (a, b) .

A proof of this test for two-variable functions is based on Taylor polynomials, and relies on symmetric matrices, eigenvalues, and quadratic forms. The steps for a proof will be found later in this section.

Introduction

We have seen how to diagonalize a matrix – if we can find n linearly independent eigenvectors of an $n \times n$ matrix A and let P be the matrix whose columns are those eigenvectors, then $P^{-1}AP$ is a diagonal matrix with the eigenvalues down the diagonal in the same order corresponding to the eigenvectors placed in P . We will see that in certain cases we can take this one step further and create an orthogonal matrix with eigenvectors as columns to diagonalize a matrix. This is called orthogonal diagonalization. Orthogonal diagonalizability is useful in that it allows us to find a “convenient” coordinate system in which to interpret the results of certain matrix transformations. A set of orthonormal basis vectors for an orthogonally diagonalizable matrix A is called a set of *principal axes* for A . Orthogonal diagonalization will also play a crucial role in the singular value decomposition of a matrix, a decomposition that has been described by some as the “pinnacle” of linear algebra.

Definition 31.3. An $n \times n$ matrix A is **orthogonally diagonalizable** if there is an orthogonal matrix P such that

$$P^T AP$$

is a diagonal matrix. We say that the matrix P *orthogonally diagonalizes* the matrix A .

Preview Activity 31.1.

- (1) For each matrix A whose eigenvalues and corresponding eigenvectors are given, find a matrix P such that $P^{-1}AP$ is a diagonal matrix.

(a) $A = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}$ with eigenvalues -1 and 3 and corresponding eigenvectors $\mathbf{v}_1 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$ and $\mathbf{v}_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$.

(b) $A = \begin{bmatrix} 1 & 2 \\ 1 & 2 \end{bmatrix}$ with eigenvalues 0 and 3 and corresponding eigenvectors $\mathbf{v}_1 = \begin{bmatrix} -2 \\ 1 \end{bmatrix}$ and $\mathbf{v}_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$.



(c) $A = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 2 \end{bmatrix}$ with eigenvalues 0, 1, and 3 and corresponding eigenvectors
 $\mathbf{v}_1 = \begin{bmatrix} -1 \\ -1 \\ 1 \end{bmatrix}$, $\mathbf{v}_2 = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}$, and $\mathbf{v}_3 = \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}$.

- (2) Which matrices in part 1 seem to satisfy the orthogonal diagonalization requirement? Do you notice any common traits among these matrices?

Symmetric Matrices

As we saw in Preview Activity 31.1, matrices that are not symmetric need not be orthogonally diagonalizable, but the symmetric matrix examples are orthogonally diagonalizable. We explore that idea in this section.

If P is a matrix that orthogonally diagonalizes the matrix A , then $P^TAP = D$, where D is a diagonal matrix. Since $D^T = D$ and $A = PDP^T$, we have

$$\begin{aligned} A &= PDP^T \\ &= PD^TP^T \\ &= (P^T)^T D^T P^T \\ &= (PDP^T)^T \\ &= A^T. \end{aligned}$$

Therefore, $A^T = A$ and matrices with this property are the only matrices that can be orthogonally diagonalized. Recall that any matrix A satisfying $A^T = A$ is a symmetric matrix.

While we have just shown that the only matrices that can be orthogonally diagonalized are the symmetric matrices, the amazing thing about symmetric matrices is that *every* symmetric matrix can be orthogonally diagonalized. We will prove this shortly.

Symmetric matrices have useful properties, a few of which are given in the following activity (we will use some of these properties later in this section).

Activity 31.1. Let A be a symmetric $n \times n$ matrix and let \mathbf{x} and \mathbf{y} be vectors in \mathbb{R}^n .

- (a) Show that $\mathbf{x}^T A \mathbf{y} = (A\mathbf{x})^T \mathbf{y}$.
- (b) Show that $(A\mathbf{x}) \cdot \mathbf{y} = \mathbf{x} \cdot (A\mathbf{y})$.
- (c) Show that the eigenvalues of a 2×2 symmetric matrix $A = \begin{bmatrix} a & b \\ b & c \end{bmatrix}$ are real.

Activity 31.1 (c) shows that a 2×2 symmetric matrix has real eigenvalues. This is a general result about real symmetric matrices.



Theorem 31.4. *Let A be an $n \times n$ symmetric matrix with real entries. Then the eigenvalues of A are real.*

Proof. Let A be an $n \times n$ symmetric matrix with real entries and let λ be an eigenvalue of A with eigenvector \mathbf{v} . To show that λ is real, we will show that $\bar{\lambda} = \lambda$. We know

$$A\mathbf{v} = \lambda\mathbf{v}. \quad (31.1)$$

Since A has real entries, we also know that $\bar{\lambda}$ is an eigenvalue for A with eigenvector $\bar{\mathbf{v}}$. Multiply both sides of (31.1) on the left by $\bar{\mathbf{v}}^T$ to obtain

$$\bar{\mathbf{v}}^T A \mathbf{v} = \bar{\mathbf{v}}^T \lambda \mathbf{v} = \lambda (\bar{\mathbf{v}}^T \mathbf{v}). \quad (31.2)$$

Now

$$\bar{\mathbf{v}}^T A \mathbf{v} = (A\bar{\mathbf{v}})^T \mathbf{v} = (\bar{\lambda} \bar{\mathbf{v}})^T \mathbf{v} = \bar{\lambda} (\bar{\mathbf{v}}^T \mathbf{v})$$

and equation (31.2) becomes

$$\bar{\lambda} (\bar{\mathbf{v}}^T \mathbf{v}) = \lambda (\bar{\mathbf{v}}^T \mathbf{v}).$$

Since $\mathbf{v} \neq 0$, this implies that $\bar{\lambda} = \lambda$ and λ is real. ■

To orthogonally diagonalize a matrix, it must be the case that eigenvectors corresponding to different eigenvalues are orthogonal. This is an important property and it would be useful to know when it happens.

Activity 31.2. Let A be a real symmetric matrix with eigenvalues λ_1 and λ_2 and corresponding eigenvectors \mathbf{v}_1 and \mathbf{v}_2 , respectively.

- (a) Use Activity 31.1 (b) to show that $\lambda_1 \mathbf{v}_1 \cdot \mathbf{v}_2 = \lambda_2 \mathbf{v}_1 \cdot \mathbf{v}_2$.
- (b) Explain why the result of part (a) shows that \mathbf{v}_1 and \mathbf{v}_2 are orthogonal if $\lambda_1 \neq \lambda_2$.

Activity 31.2 proves the following theorem.

Theorem 31.5. *If A is a real symmetric matrix, then eigenvectors corresponding to distinct eigenvalues are orthogonal.*

Recall that the only matrices than can be orthogonally diagonalized are the symmetric matrices. Now we show that every real symmetric matrix can be orthogonally diagonalized, which completely characterizes the matrices that are orthogonally diagonalizable.

Theorem 31.6. *Let A be a real symmetric matrix. Then A is orthogonally diagonalizable.*

Proof. We will assume that all matrices are real matrices. To prove that every symmetric matrix is orthogonally diagonalizable, we will proceed by contradiction and assume that there are $n \times n$ symmetric matrices that are not orthogonally diagonalizable for some values of n . Since n must be positive (greater than 1, in fact, since every 1×1 matrix is orthogonally diagonalizable), there must be a smallest value of n so that there is an $n \times n$ symmetric matrix that is not orthogonally diagonalizable. Let A be one of these smallest $n \times n$ matrices that is not orthogonally diagonalizable. By



our assumption, every $k \times k$ symmetric matrix M with $k < n$ is orthogonally diagonalizable. Since A is not orthogonally diagonalizable, there cannot exist an orthogonal basis for \mathbb{R}^n that consists of eigenvectors of A . We will now show that this is impossible, which will force our assumption that A is not orthogonally diagonalizable to be false.

Since A is a symmetric matrix, we know that A has n (counting multiplicities) real eigenvalues (not necessarily distinct). Let λ_1 be one of these real eigenvalues and \mathbf{u}_1 a corresponding unit eigenvector. Let

$$W = \text{Span}\{\mathbf{u}_1\}^\perp.$$

So $\mathbf{x} \in W$ if $0 = \mathbf{u}_1 \cdot \mathbf{x} = \mathbf{u}_1^\top \mathbf{x}$. Now W is a subspace of \mathbb{R}^n and $\text{Span}\{\mathbf{u}_1\}$ has dimension 1, so $\dim(W) = n - 1$. We can then construct an orthonormal basis $\mathcal{B} = \{\mathbf{u}_2, \mathbf{u}_3, \dots, \mathbf{u}_n\}$ for W .

Note that W is invariant under left multiplication by A . To see why, let $\mathbf{w} \in W$. Then

$$\mathbf{u}_1 \cdot A\mathbf{w} = (A\mathbf{u}_1)^\top \mathbf{w} = \lambda_1 \mathbf{u}_1^\top \mathbf{w} = \lambda_1 \mathbf{u}_1 \cdot \mathbf{w} = 0,$$

so $A\mathbf{w}$ is orthogonal to \mathbf{u}_1 and is therefore in W .

Since W is invariant under left multiplication by A , we can think of left multiplication by A as a linear transformation from the $n - 1$ dimensional vector space W to itself. We call this the *restriction* of A to W and will denote it by $A|_W$. The coordinate transformation T defined by $T(\mathbf{x}) = [\mathbf{x}]_{\mathcal{B}}$ maps the $n - 1$ dimensional vector space W to \mathbb{R}^{n-1} and we know that T is an invertible map. We can then view left multiplication by A on W as a composite of the coordinate map T , a matrix transformation M from \mathbb{R}^{n-1} to \mathbb{R}^{n-1} and then T^{-1} as shown in Figure 31.1.

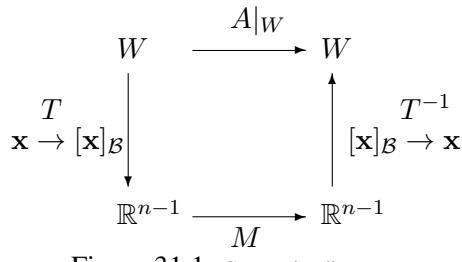


Figure 31.1: Composite diagram

Let us focus on the coordinate mapping T for a moment. This mapping has two important properties that we will need.

Claim.

- (1) $T(\mathbf{x}) \cdot T(\mathbf{y}) = \mathbf{x} \cdot \mathbf{y}$ for every \mathbf{x}, \mathbf{y} in W .
- (2) $T^{-1}([\mathbf{x}]_{\mathcal{B}}) \cdot T^{-1}([\mathbf{y}]_{\mathcal{B}}) = [\mathbf{x}]_{\mathcal{B}} \cdot [\mathbf{y}]_{\mathcal{B}}$ for every $[\mathbf{x}]_{\mathcal{B}}, [\mathbf{y}]_{\mathcal{B}}$ in \mathbb{R}^{n-1} .

Proof of the Claim: Let \mathbf{x}, \mathbf{y} in W . Since \mathcal{B} is a basis for W , we know that

$$\mathbf{x} = x_2 \mathbf{u}_2 + x_3 \mathbf{u}_3 + \cdots + x_n \mathbf{u}_n \text{ and } \mathbf{y} = y_2 \mathbf{u}_2 + y_3 \mathbf{u}_3 + \cdots + y_n \mathbf{u}_n$$

for some scalars x_2, x_3, \dots, x_n and y_2, y_3, \dots, y_n . So $[\mathbf{x}]_{\mathcal{B}} = [x_2 \ x_3 \ \dots \ x_n]^\top$ and $[\mathbf{y}]_{\mathcal{B}} = [y_2 \ y_3 \ \dots \ y_n]^\top$ and it follows that

$$T(\mathbf{x}) \cdot T(\mathbf{y}) = [\mathbf{x}]_{\mathcal{B}} \cdot [\mathbf{y}]_{\mathcal{B}} = x_2 y_2 + x_3 y_3 + \cdots + x_n y_n. \quad (31.3)$$

Now \mathcal{B} is an orthonormal basis, so $\mathbf{u}_i \cdot \mathbf{u}_j = 0$ for $i \neq j$, $\mathbf{u}_i \cdot \mathbf{u}_i = 1$ for all i , and

$$\begin{aligned}\mathbf{x} \cdot \mathbf{y} &= (x_2\mathbf{u}_2 + x_3\mathbf{u}_3 + \cdots + x_n\mathbf{u}_n) \cdot (y_2\mathbf{u}_2 + y_3\mathbf{u}_3 + \cdots + y_n\mathbf{u}_n) \\ &= (x_2y_2\mathbf{u}_2 \cdot \mathbf{u}_2) + (x_3y_3\mathbf{u}_3 \cdot \mathbf{u}_3) + \cdots + (x_ny_n\mathbf{u}_n \cdot \mathbf{u}_n) \\ &= x_2y_2 + x_3y_3 + \cdots + x_ny_n.\end{aligned}\tag{31.4}$$

Thus, $T(\mathbf{x}) \cdot T(\mathbf{y}) = \mathbf{x} \cdot \mathbf{y}$, proving part 1 of the claim.

For part 2, note that

$$T^{-1}([\mathbf{x}]_{\mathcal{B}}) \cdot T^{-1}([\mathbf{y}]_{\mathcal{B}}) = \mathbf{x} \cdot \mathbf{y} = [\mathbf{x}]_{\mathcal{B}} \cdot [\mathbf{y}]_{\mathcal{B}}$$

by (31.3) and (31.4), verifying part 2 of the claim. ■

We now apply this claim to show that the matrix M is a symmetric matrix. Since the coordinate transformation T is an onto mapping, every vector in \mathbb{R}^{n-1} can be written as $[\mathbf{x}]_{\mathcal{B}}$ for some \mathbf{x} in W . Let $[\mathbf{x}]_{\mathcal{B}}$ and $[\mathbf{y}]_{\mathcal{B}}$ be arbitrary vectors in \mathbb{R}^{n-1} . From Figure 31.1 we can see that $M = TA|_W T^{-1} = TAT^{-1}$, so

$$\begin{aligned}M[\mathbf{x}]_{\mathcal{B}} \cdot [\mathbf{y}]_{\mathcal{B}} &= TAT^{-1}([\mathbf{x}]_{\mathcal{B}}) \cdot [\mathbf{y}]_{\mathcal{B}} \\ &= TAT^{-1}([\mathbf{x}]_{\mathcal{B}}) \cdot TT^{-1}([\mathbf{y}]_{\mathcal{B}}) \\ &= AT^{-1}([\mathbf{x}]_{\mathcal{B}}) \cdot T^{-1}([\mathbf{y}]_{\mathcal{B}}) \\ &= T^{-1}([\mathbf{x}]_{\mathcal{B}}) \cdot AT^{-1}([\mathbf{y}]_{\mathcal{B}}) \\ &= TT^{-1}([\mathbf{x}]_{\mathcal{B}}) \cdot TAT^{-1}([\mathbf{y}]_{\mathcal{B}}) \\ &= [\mathbf{x}]_{\mathcal{B}} \cdot M[\mathbf{y}]_{\mathcal{B}}.\end{aligned}$$

Therefore, M is a symmetric matrix. Since M is $(n-1) \times (n-1)$, we can conclude that M is orthogonally diagonalizable. Thus, there is an orthonormal basis $\{[\mathbf{x}_2]_{\mathcal{B}}, [\mathbf{x}_3]_{\mathcal{B}}, \dots, [\mathbf{x}_n]_{\mathcal{B}}\}$ for \mathbb{R}^{n-1} consisting of eigenvectors of M . Let $M[\mathbf{x}_i]_{\mathcal{B}} = \lambda_i[\mathbf{x}_i]_{\mathcal{B}}$ for each $2 \leq i \leq n$.

Finally, we will show that the set $\mathcal{C} = \{\mathbf{u}_1, \mathbf{x}_2, \mathbf{x}_3, \dots, \mathbf{x}_n\}$ is an orthonormal set of eigenvectors of A . This will contradict the fact that we assumed that A was *not* orthogonally diagonalizable. The set \mathcal{B} is an orthonormal subset of W and so by the definition of W each vector in \mathcal{B} is orthogonal to \mathbf{u}_1 . Thus the set \mathcal{C} is an orthogonal set. To complete our proof we only need show that each \mathbf{x}_i is an eigenvector of A . Choose an i arbitrarily between 2 and n . Now

$$A\mathbf{x}_i = T^{-1}MT(\mathbf{x}_i) = T^{-1}M[\mathbf{x}_i]_{\mathcal{B}} = T^{-1}(\lambda_i[\mathbf{x}_i]_{\mathcal{B}}) = \lambda_iT^{-1}([\mathbf{x}_i]_{\mathcal{B}}) = \lambda_i\mathbf{x}_i.$$

Therefore, the set $\mathcal{C} = \{\mathbf{u}_1, \mathbf{x}_2, \mathbf{x}_3, \dots, \mathbf{x}_n\}$ is an orthogonal set of eigenvectors of A . Since the transformations T and T^{-1} preserve dot products, they also preserve lengths. So, in fact, \mathcal{C} is an orthonormal set of eigenvectors of A as desired. This completes our proof. ■

The set of eigenvalues of a matrix A is called the *spectrum* of A and we have just proved the following theorem.

Theorem 31.7 (The Spectral Theorem for Real Symmetric Matrices). *Let A be an $n \times n$ symmetric matrix with real entries. Then*

- (1) *A has n real eigenvalues (counting multiplicities)*



- (2) the dimension of each eigenspace of A is the multiplicity of the corresponding eigenvalue as a root of the characteristic polynomial
- (3) eigenvectors corresponding to different eigenvalues are orthogonal
- (4) A is orthogonally diagonalizable.

So any real symmetric matrix is orthogonally diagonalizable. We have seen examples of the orthogonal diagonalization of $n \times n$ real symmetric matrices with n distinct eigenvalues, but how do we orthogonally diagonalize a symmetric matrix having eigenvalues of multiplicity greater than 1? The next activity shows us the process.

Activity 31.3. Let $A = \begin{bmatrix} 4 & 2 & 2 \\ 2 & 4 & 2 \\ 2 & 2 & 4 \end{bmatrix}$. The eigenvalues of A are 2 and 8, with eigenspace of dimension 2 and dimension 1, respectively.

- (a) Explain why A can be orthogonally diagonalized.
- (b) Two linearly independent eigenvectors for A corresponding to the eigenvalue 2 are $\mathbf{v}_1 = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$ and $\mathbf{v}_2 = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}$. Note that $\mathbf{v}_1, \mathbf{v}_2$ are not orthogonal, so cannot be in an orthogonal basis of \mathbb{R}^3 consisting of eigenvectors of A . So find a set $\{\mathbf{w}_1, \mathbf{w}_2\}$ of orthogonal eigenvectors of A so that $\text{Span}\{\mathbf{w}_1, \mathbf{w}_2\} = \text{Span}\{\mathbf{v}_1, \mathbf{v}_2\}$.
- (c) The vector $\mathbf{v}_3 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ is an eigenvector for A corresponding to the eigenvalue 8. What can you say about the orthogonality relationship between \mathbf{w}_i 's and \mathbf{v}_3 ?
- (d) Find a matrix P that orthogonally diagonalizes A . Verify your work.

The Spectral Decomposition of a Symmetric Matrix A

Let A be an $n \times n$ symmetric matrix with real entries. The Spectral Theorem tells us we can find an orthonormal basis $\{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n\}$ of eigenvectors of A . Let $A\mathbf{u}_i = \lambda_i \mathbf{u}_i$ for each $1 \leq i \leq n$. If $P = [\mathbf{u}_1 \ \mathbf{u}_2 \ \mathbf{u}_3 \ \cdots \ \mathbf{u}_n]$, then we know that

$$P^T A P = P^{-1} A P = D,$$

where D is the $n \times n$ diagonal matrix

$$\begin{bmatrix} \lambda_1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & \lambda_2 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & \lambda_n \end{bmatrix}.$$



Since $A = PDP^T$ we see that

$$\begin{aligned}
 A &= [\mathbf{u}_1 \mathbf{u}_2 \mathbf{u}_3 \cdots \mathbf{u}_n] \begin{bmatrix} \lambda_1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & \lambda_2 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & \lambda_n \end{bmatrix} \begin{bmatrix} \mathbf{u}_1^T \\ \mathbf{u}_2^T \\ \mathbf{u}_3^T \\ \vdots \\ \mathbf{u}_n^T \end{bmatrix} \\
 &= [\lambda_1 \mathbf{u}_1 \ \lambda_2 \mathbf{u}_2 \ \lambda_3 \mathbf{u}_3 \ \cdots \ \lambda_n \mathbf{u}_n] \begin{bmatrix} \mathbf{u}_1^T \\ \mathbf{u}_2^T \\ \mathbf{u}_3^T \\ \vdots \\ \mathbf{u}_n^T \end{bmatrix} \\
 &= \lambda_1 \mathbf{u}_1 \mathbf{u}_1^T + \lambda_2 \mathbf{u}_2 \mathbf{u}_2^T + \lambda_3 \mathbf{u}_3 \mathbf{u}_3^T + \cdots + \lambda_n \mathbf{u}_n \mathbf{u}_n^T, \tag{31.5}
 \end{aligned}$$

where the last product follows from Exercise 4. The expression in (31.5) is called a *spectral decomposition* of the matrix A . Let $P_i = \mathbf{u}_i \mathbf{u}_i^T$ for each i . The matrices P_i satisfy several special conditions given in the next theorem. The proofs are left to the exercises.

Theorem 31.8. *Let A be an $n \times n$ symmetric matrix with real entries, and let $\{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n\}$ be an orthonormal basis of eigenvectors of A with $A\mathbf{u}_i = \lambda_i \mathbf{u}_i$ for each i . For each i , let $P_i = \mathbf{u}_i \mathbf{u}_i^T$. Then*

- (1) $A = \lambda_1 P_1 + \lambda_2 P_2 + \cdots + \lambda_n P_n$,
- (2) P_i is a symmetric matrix for each i ,
- (3) P_i is a rank 1 matrix for each i ,
- (4) $P_i^2 = P_i$ for each i ,
- (5) $P_i P_j = 0$ if $i \neq j$,
- (6) $P_i \mathbf{u}_i = \mathbf{u}_i$ for each i ,
- (7) $P_i \mathbf{u}_j = \mathbf{0}$ if $i \neq j$,
- (8) For any vector \mathbf{v} in \mathbb{R}^n , $P_i \mathbf{v} = \text{proj}_{\text{Span}\{\mathbf{u}_i\}} \mathbf{v}$.

The consequence of Theorem 31.8 is that any symmetric matrix can be written as the sum of symmetric, rank 1 matrices. As we will see later, this kind of decomposition contains much information about the matrix product $A^T A$ for any matrix A .

Activity 31.4. Let $A = \begin{bmatrix} 4 & 2 & 2 \\ 2 & 4 & 2 \\ 2 & 2 & 4 \end{bmatrix}$. Let $\lambda_1 = 2$, $\lambda_2 = 2$, and $\lambda_3 = 8$ be the eigenvalues of A .

A basis for the eigenspace E_8 of A corresponding to the eigenvalue 8 is $\{[1 \ 1 \ 1]^T\}$ and a basis for the eigenspace E_2 of A corresponding to the eigenvalue 2 is $\{[1 \ -1 \ 0]^T, [1 \ 0 \ -1]^T\}$. (Compare to Activity 31.3.)



- (a) Find orthonormal eigenvectors \mathbf{u}_1 , \mathbf{u}_2 , and \mathbf{u}_3 of A corresponding to λ_1 , λ_2 , and λ_3 , respectively.
- (b) Compute $\lambda_1 \mathbf{u}_1 \mathbf{u}_1^\top$
- (c) Compute $\lambda_2 \mathbf{u}_2 \mathbf{u}_2^\top$
- (d) Compute $\lambda_3 \mathbf{u}_3 \mathbf{u}_3^\top$
- (e) Verify that $A = \lambda_1 \mathbf{u}_1 \mathbf{u}_1^\top + \lambda_2 \mathbf{u}_2 \mathbf{u}_2^\top + \lambda_3 \mathbf{u}_3 \mathbf{u}_3^\top$.

Examples

What follows are worked examples that use the concepts from this section.

Example 31.9. For each of the following matrices A , determine if A is diagonalizable. If A is not diagonalizable, explain why. If A is diagonalizable, find a matrix P so that $P^{-1}AP$ is a diagonal matrix. If the matrix is diagonalizable, is it orthogonally diagonalizable? If orthogonally diagonalizable, find an orthogonal matrix that diagonalizes A . Use appropriate technology to find eigenvalues and eigenvectors.

$$(a) A = \begin{bmatrix} 2 & 0 & 0 \\ -1 & 3 & 2 \\ 1 & -1 & 0 \end{bmatrix} \quad (b) A = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad (c) A = \begin{bmatrix} 4 & 2 & 1 \\ 2 & 7 & 2 \\ 1 & 2 & 4 \end{bmatrix}$$

Example Solution.

- (a) Recall that an $n \times n$ matrix A is diagonalizable if and only if A has n linearly independent eigenvectors, and A is orthogonally diagonalizable if and only if A is symmetric. Since A is not symmetric, A is not orthogonally diagonalizable. Technology shows that the eigenvalues of A are 2 and 1 and bases for the corresponding eigenspaces are $\{[1 \ 1 \ 0]^\top, [2 \ 0 \ 1]^\top\}$

and $\{[0 \ -1 \ 1]^\top\}$. So A is diagonalizable and if $P = \begin{bmatrix} 1 & 2 & 0 \\ 1 & 0 & -1 \\ 0 & 1 & 1 \end{bmatrix}$, then

$$P^{-1}AP = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

- (b) Since A is not symmetric, A is not orthogonally diagonalizable. Technology shows that the eigenvalues of A are 0 and 1 and bases for the corresponding eigenspaces are $\{[0 \ 0 \ 1]^\top\}$ and $\{[1 \ 0 \ 0]^\top\}$. We cannot create a basis of \mathbb{R}^3 consisting of eigenvectors of A , so A is not diagonalizable.
- (c) Since A is symmetric, A is orthogonally diagonalizable. Technology shows that the eigenvalues of A are 3 and 9 and bases for the eigenspaces $\{[-1 \ 0 \ 1]^\top, [-2 \ 1 \ 0]^\top\}$ and $\{[1 \ 2 \ 1]^\top\}$, respectively. To find an orthogonal matrix that diagonalizes A , we must find an orthonormal basis of \mathbb{R}^3 consisting of eigenvectors of A . To do that, we use the Gram-Schmidt process



to obtain an orthogonal basis for the eigenspace of A corresponding to the eigenvalue 3. Doing so gives an orthogonal basis $\{\mathbf{v}_1, \mathbf{v}_2\}$, where $\mathbf{v}_1 = [-1 \ 0 \ 1]^\top$ and

$$\begin{aligned}\mathbf{v}_2 &= [-2 \ 1 \ 0]^\top - \frac{[-2 \ 1 \ 0]^\top \cdot [-1 \ 0 \ 1]^\top}{[-1 \ 0 \ 1]^\top \cdot [-1 \ 0 \ 1]^\top} [-1 \ 0 \ 1]^\top \\ &= [-2 \ 1 \ 0]^\top - [-1 \ 0 \ 1]^\top \\ &= [-1 \ 1 \ -1]^\top.\end{aligned}$$

So an orthonormal basis for \mathbb{R}^3 of eigenvectors of A is

$$\left\{ \frac{1}{\sqrt{2}}[-1 \ 0 \ 1]^\top, \frac{1}{\sqrt{3}}[-1 \ 1 \ -1]^\top, \frac{1}{\sqrt{6}}[1 \ 1 \ 1]^\top \right\}.$$

Therefore, A is orthogonally diagonalizable and if P is the matrix $\begin{bmatrix} -\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{3}} & \frac{1}{\sqrt{6}} \\ 0 & \frac{1}{\sqrt{3}} & \frac{2}{\sqrt{6}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{3}} & \frac{1}{\sqrt{6}} \end{bmatrix}$, then

$$P^{-1}AP = \begin{bmatrix} 3 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 9 \end{bmatrix}.$$

Example 31.10. Let $A = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}$. Find an orthonormal basis for \mathbb{R}^4 consisting of eigenvectors of A .

Example Solution.

Since A is symmetric, there is an orthogonal matrix P such that $P^{-1}AP$ is diagonal. The columns of P will form an orthonormal basis for \mathbb{R}^4 . Using a cofactor expansion along the first row shows that

$$\begin{aligned}\det(A - \lambda I_4) &= \det \left(\begin{bmatrix} -\lambda & 0 & 0 & 1 \\ 0 & -\lambda & 1 & 0 \\ 0 & 1 & -\lambda & 0 \\ 1 & 0 & 0 & -\lambda \end{bmatrix} \right) \\ &= (\lambda^2 - 1)^2 \\ &= (\lambda + 1)^2(\lambda - 1)^2.\end{aligned}$$

So the eigenvalues of A are 1 and -1 . The reduced row echelon forms of $A - I_4$ and $A + I_4$ are, respectively,

$$\begin{bmatrix} 1 & 0 & 0 & -1 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \text{ and } \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

Thus, a basis for the eigenspace E_1 of A is $\{[0 \ 1 \ 1 \ 0]^\top, [1 \ 0 \ 0 \ 1]^\top\}$ and a basis for the eigenspace E_{-1} of A is $\{[0 \ 1 \ -1 \ 0]^\top, [1 \ 0 \ 0 \ -1]^\top\}$. The set $\{[0 \ 1 \ 1 \ 0]^\top, [1 \ 0 \ 0 \ 1]^\top, [0 \ 1 \ -1 \ 0]^\top, [1 \ 0 \ 0 \ -1]^\top\}$



is an orthogonal set, so an orthonormal basis for \mathbb{R}^4 consisting of eigenvectors of A is

$$\left\{ \frac{1}{\sqrt{2}}[0 \ 1 \ 1 \ 0]^\top, \frac{1}{\sqrt{2}}[1 \ 0 \ 0 \ 1]^\top, \frac{1}{\sqrt{2}}[0 \ 1 \ -1 \ 0]^\top, \frac{1}{\sqrt{2}}[1 \ 0 \ 0 \ -1]^\top \right\}.$$

Summary

- An $n \times n$ matrix A is orthogonally diagonalizable if there is an orthogonal matrix P such that $P^\top AP$ is a diagonal matrix. Orthogonal diagonalizability is useful in that it allows us to find a “convenient” coordinate system in which to interpret the results of certain matrix transformations. Orthogonal diagonalization also plays a crucial role in the singular value decomposition of a matrix.
- An $n \times n$ matrix A is symmetric if $A^\top = A$. The symmetric matrices are exactly the matrices that can be orthogonally diagonalized.
- The spectrum of a matrix is the set of eigenvalues of the matrix.

Exercises

- (1) For each of the following matrices, find an orthogonal matrix P so that $P^\top AP$ is a diagonal matrix, or explain why no such matrix exists.

$$(a) A = \begin{bmatrix} 3 & -4 \\ -4 & -3 \end{bmatrix} \quad (b) A = \begin{bmatrix} 4 & 1 & 1 \\ 1 & 1 & 4 \\ 1 & 4 & 1 \end{bmatrix} \quad (c) A = \begin{bmatrix} 1 & 2 & 0 & 0 \\ 0 & 1 & 2 & 1 \\ 1 & 1 & 1 & 1 \\ 3 & 0 & 5 & 2 \end{bmatrix}$$

- (2) For each of the following matrices find an orthonormal basis of eigenvectors of A . Then find a spectral decomposition of A .

$$(a) A = \begin{bmatrix} 3 & -4 \\ -4 & -3 \end{bmatrix} \quad (b) A = \begin{bmatrix} 4 & 1 & 1 \\ 1 & 1 & 4 \\ 1 & 4 & 1 \end{bmatrix}$$

$$(c) A = \begin{bmatrix} -4 & 0 & -24 \\ 0 & -8 & 0 \\ -24 & 0 & 16 \end{bmatrix} \quad (d) A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 2 \\ 0 & 2 & -3 \end{bmatrix}$$

- (3) Find a non-diagonal 4×4 matrix with eigenvalues 2, 3 and 6 which can be orthogonally diagonalized.

- (4) Let $A = [a_{ij}] = [\mathbf{c}_1 \ \mathbf{c}_2 \ \cdots \ \mathbf{c}_m]$ be an $k \times m$ matrix with columns $\mathbf{c}_1, \mathbf{c}_2, \dots, \mathbf{c}_m$, and let



$B = [b_{ij}] = \begin{bmatrix} \mathbf{r}_1 \\ \mathbf{r}_2 \\ \vdots \\ \mathbf{r}_m \end{bmatrix}$ be an $m \times n$ matrix with rows $\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_m$. Show that

$$AB = [\mathbf{c}_1 \ \mathbf{c}_2 \ \cdots \ \mathbf{c}_m] \begin{bmatrix} \mathbf{r}_1 \\ \mathbf{r}_2 \\ \vdots \\ \mathbf{r}_m \end{bmatrix} = \mathbf{c}_1\mathbf{r}_1 + \mathbf{c}_2\mathbf{r}_2 + \cdots + \mathbf{c}_m\mathbf{r}_m.$$

- (5) Let A be an $n \times n$ symmetric matrix with real entries and let $\{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n\}$ be an orthonormal basis of eigenvectors of A . For each i , let $P_i = \mathbf{u}_i\mathbf{u}_i^\top$. Prove Theorem 31.8 – that is, verify each of the following statements.

- (a) For each i , P_i is a symmetric matrix.
- (b) For each i , P_i is a rank 1 matrix.
- (c) For each i , $P_i^2 = P_i$.
- (d) If $i \neq j$, then $P_iP_j = 0$.
- (e) For each i , $P_i\mathbf{u}_i = \mathbf{u}_i$.
- (f) If $i \neq j$, then $P_i\mathbf{u}_j = 0$.
- (g) If \mathbf{v} is in \mathbb{R}^n , show that

$$P_i\mathbf{v} = \text{proj}_{\text{Span}\{\mathbf{u}_i\}}\mathbf{v}.$$

For this reason we call P_i an *orthogonal projection matrix*.

- (6) Show that if M is an $n \times n$ matrix and $(M\mathbf{x}) \cdot \mathbf{y} = \mathbf{x} \cdot (M\mathbf{y})$ for every \mathbf{x}, \mathbf{y} in \mathbb{R}^n , then M is a symmetric matrix. (Hint: Try $\mathbf{x} = \mathbf{e}_i$ and $\mathbf{y} = \mathbf{e}_j$.)
- (7) Let A be an $n \times n$ symmetric matrix and assume that A has an orthonormal basis $\{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n\}$ of eigenvectors of A so that $A\mathbf{u}_i = \lambda_i\mathbf{u}_i$ for each i . Let $P_i = \mathbf{u}_i\mathbf{u}_i^\top$ for each i . It is possible that not all of the eigenvalues of A are distinct. In this case, some of the eigenvalues will be repeated in the spectral decomposition of A . If we want only distinct eigenvalues to appear, we might do the following. Let $\mu_1, \mu_2, \dots, \mu_k$ be the distinct eigenvalues of A . For each j between 1 and k , let Q_j be the sum of all of the P_i that have μ_j as eigenvalue.

- (a) The eigenvalues for the matrix $A = \begin{bmatrix} 0 & 2 & 0 & 0 \\ 2 & 3 & 0 & 0 \\ 0 & 0 & 0 & 2 \\ 0 & 0 & 2 & 3 \end{bmatrix}$ are -1 and 4 . Find a basis for

each eigenspace and determine each P_i . Then find k, μ_1, \dots, μ_k , and each Q_j .

- (b) Show in general (not just for the specific example in part (a), that the Q_j satisfy the same properties as the P_i . That is, verify the following.
- i. $A = \mu_1Q_1 + \mu_2Q_2 + \cdots + \mu_kQ_k$
 - ii. Q_j is a symmetric matrix for each j
 - iii. $Q_j^2 = Q_j$ for each j



- iv. $Q_j Q_\ell = 0$ when $j \neq \ell$
- v. if E_{μ_j} is the eigenspace for A corresponding to the eigenvalue μ_j , and if \mathbf{v} is in \mathbb{R}^n , then $Q_j \mathbf{v} = \text{proj}_{E_{\mu_j}} \mathbf{v}$.
- (c) What is the rank of Q_j ? Verify your answer.
- (8) Label each of the following statements as True or False. Provide justification for your response.
- True/False** Every real symmetric matrix is diagonalizable.
 - True/False** If P is a matrix whose columns are eigenvectors of a symmetric matrix, then the columns of P are orthogonal.
 - True/False** If A is a symmetric matrix, then eigenvectors of A corresponding to distinct eigenvalues are orthogonal.
 - True/False** If \mathbf{v}_1 and \mathbf{v}_2 are distinct eigenvectors of a symmetric matrix A , then \mathbf{v}_1 and \mathbf{v}_2 are orthogonal.
 - True/False** Any symmetric matrix can be written as a sum of symmetric rank 1 matrices.
 - True/False** If A is a matrix satisfying $A^\top = A$, and \mathbf{u} and \mathbf{v} are vectors satisfying $A\mathbf{u} = 2\mathbf{u}$ and $A\mathbf{v} = -2\mathbf{v}$, then $\mathbf{u} \cdot \mathbf{v} = 0$.
 - True/False** If an $n \times n$ matrix A has n orthogonal eigenvectors, then A is a symmetric matrix.
 - True/False** If an $n \times n$ matrix has n real eigenvalues (counted with multiplicity), then A is a symmetric matrix.
 - True/False** For each eigenvalue of a symmetric matrix, the algebraic multiplicity equals the geometric multiplicity.
 - True/False** If A is invertible and orthogonally diagonalizable, then so is A^{-1} .
 - True/False** If A, B are orthogonally diagonalizable $n \times n$ matrices, then so is AB .

Project: The Second Derivative Test for Functions of Two Variables

In this project we will verify the Second Derivative Test for functions of two variables.¹ This test will involve Taylor polynomials and linear algebra. As a quick review, recall that the second order Taylor polynomial for a function f of a single variable x at $x = a$ is

$$P_2(x) = f(a) + f'(a)(x - a) + \frac{f''(a)}{2}(x - a)^2. \quad (31.6)$$

As with the linearization of a function, the second order Taylor polynomial is a good approximation to f around a – that is $f(x) \approx P_2(x)$ for x close to a . If a is a critical number for f with $f'(a) = 0$,

¹Many thanks to Professor Paul Fishback for sharing his activity on this topic. Much of this project comes from his activity.



then

$$P_2(x) = f(a) + \frac{f''(a)}{2}(x - a)^2.$$

In this situation, if $f''(a) < 0$, then $\frac{f''(a)}{2}(x - a)^2 \leq 0$ for x close to a , which makes $P_2(x) \leq f(a)$. This implies that $f(x) \approx P_2(x) \leq f(a)$ for x close to a , which makes $f(a)$ a relative maximum value for f . Similarly, if $f''(a) > 0$, then $f(a)$ is a relative minimum.

We now need a Taylor polynomial for a function of two variables. The complication of the additional independent variable in the two variable case means that the Taylor polynomials will need to contain all of the possible monomials of the indicated degrees. Recall that the linearization (or tangent plane) to a function $f = f(x, y)$ at a point (a, b) is given by

$$P_1(x, y) = f(a, b) + f_x(a, b)(x - a) + f_y(a, b)(y - b).$$

Note that $P_1(a, b) = f(a, b)$, $\frac{\partial P_1}{\partial x}(a, b) = f_x(a, b)$, and $\frac{\partial P_1}{\partial y}(a, b) = f_y(a, b)$. This makes $P_1(x, y)$ the best linear approximation to f near the point (a, b) . The polynomial $P_1(x, y)$ is the first order Taylor polynomial for f at (a, b) .

Similarly, the second order Taylor polynomial $P_2(x, y)$ centered at the point (a, b) for the function f is

$$\begin{aligned} P_2(x, y) &= f(a, b) + f_x(a, b)(x - a) + f_y(a, b)(y - b) + \frac{f_{xx}(a, b)}{2}(x - a)^2 \\ &\quad + f_{xy}(a, b)(x - a)(y - b) + \frac{f_{yy}(a, b)}{2}(y - b)^2. \end{aligned}$$

Project Activity 31.1. To see that $P_2(x, y)$ is the best approximation for f near (a, b) , we need to know that the first and second order partial derivatives of P_2 agree with the corresponding partial derivatives of f at the point (a, b) . Verify that this is true.

We can rewrite this second order Taylor polynomial using matrices and vectors so that we can apply techniques from linear algebra to analyze it. Note that

$$\begin{aligned} P_2(x, y) &= f(a, b) + \nabla f(a, b)^T \begin{bmatrix} x - a \\ y - b \end{bmatrix} \\ &\quad + \frac{1}{2} \begin{bmatrix} x - a \\ y - b \end{bmatrix}^T \begin{bmatrix} f_{xx}(a, b) & f_{xy}(a, b) \\ f_{xy}(a, b) & f_{yy}(a, b) \end{bmatrix} \begin{bmatrix} x - a \\ y - b \end{bmatrix}, \end{aligned} \quad (31.7)$$

where $\nabla f(x, y) = \begin{bmatrix} f_x(x, y) \\ f_y(x, y) \end{bmatrix}$ is the gradient of f and H is the *Hessian* of f , where $H(x, y) = \begin{bmatrix} f_{xx}(x, y) & f_{xy}(x, y) \\ f_{yx}(x, y) & f_{yy}(x, y) \end{bmatrix}$.²

Project Activity 31.2. Use Equation (31.7) to compute $P_2(x, y)$ for $f(x, y) = x^4 + y^4 - 4xy + 1$ at $(a, b) = (2, 3)$.

²Note that under reasonable conditions (e.g., that f has continuous second order mixed partial derivatives in some open neighborhood containing (x, y)) we have that $f_{xy}(x, y) = f_{yx}(x, y)$ and $H(x, y) = \begin{bmatrix} f_{xx}(a, b) & f_{xy}(a, b) \\ f_{xy}(a, b) & f_{yy}(a, b) \end{bmatrix}$ is a symmetric matrix. We will only consider functions that satisfy these reasonable conditions.



The important idea for us is that if (a, b) is a point at which f_x and f_y are zero, then ∇f is the zero vector and Equation (31.7) reduces to

$$P_2(x, y) = f(a, b) + \frac{1}{2} \begin{bmatrix} x - a \\ y - b \end{bmatrix}^T \begin{bmatrix} f_{xx}(a, b) & f_{xy}(a, b) \\ f_{xy}(a, b) & f_{yy}(a, b) \end{bmatrix} \begin{bmatrix} x - a \\ y - b \end{bmatrix}, \quad (31.8)$$

To make the connection between the multivariable second derivative test and properties of the Hessian, $H(a, b)$, at a critical point of a function f at which $\nabla f = \mathbf{0}$, we will need to connect the eigenvalues of a matrix to the determinant and the trace.

Let A be an $n \times n$ matrix with eigenvalues $\lambda_1, \lambda_2, \dots, \lambda_n$ (not necessarily distinct). Exercise 2 in Section 17 shows that

$$\det(A) = \lambda_1 \lambda_2 \cdots \lambda_n. \quad (31.9)$$

In other words, the determinant of a matrix is equal to the product of the eigenvalues of the matrix. In addition, Exercise 9 in Section 18 shows that

$$\text{trace}(A) = \lambda_1 + \lambda_2 + \cdots + \lambda_n. \quad (31.10)$$

for a diagonalizable matrix, where $\text{trace}(A)$ is the sum of the diagonal entries of A . Equation (31.10) is true for any square matrix, but we don't need the more general result for this project.

The fact that the Hessian is a symmetric matrix makes it orthogonally diagonalizable. We denote the eigenvalues of $H(a, b)$ as λ_1 and λ_2 . Thus there exists an orthogonal matrix P and a diagonal matrix $D = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}$ such that $P^T H(a, b) P = D$, or $H(a, b) = P D P^T$. Equations 31.9 and 31.10 show that

$$\lambda_1 \lambda_2 = f_{xx}(a, b) f_{yy}(a, b) - f_{xy}(a, b)^2 \text{ and } \lambda_1 + \lambda_2 = f_{xx}(a, b) + f_{yy}(a, b).$$

Now we have the machinery to verify the Second Derivative Test for Two-Variable Functions. We assume (a, b) is a point in the domain of a function f so that $\nabla f(a, b) = \mathbf{0}$. First we consider the case where $f_{xx}(a, b) f_{yy}(a, b) - f_{xy}(a, b)^2 < 0$.

Project Activity 31.3. Explain why if $f_{xx}(a, b) f_{yy}(a, b) - f_{xy}(a, b)^2 < 0$, then

$$\begin{bmatrix} x - a \\ y - b \end{bmatrix}^T H(a, b) \begin{bmatrix} x - a \\ y - b \end{bmatrix}$$

is indefinite. Explain why this implies that f is “saddle-shaped” near (a, b) . (Hint: Substitute $\mathbf{w} = \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} = P^T \begin{bmatrix} x - a \\ y - b \end{bmatrix}$. What does the graph of f look like in the w_1 and w_2 directions?)

Now we examine the situation when $f_{xx}(a, b) f_{yy}(a, b) - f_{xy}(a, b)^2 > 0$.

Project Activity 31.4. Assume that $f_{xx}(a, b) f_{yy}(a, b) - f_{xy}(a, b)^2 > 0$.

- (a) Explain why either both $f_{xx}(a, b)$ and $f_{yy}(a, b)$ are positive or both are negative.
- (b) If $f_{xx}(a, b) > 0$ and $f_{yy}(a, b) > 0$, explain why λ_1 and λ_2 must be positive.



- (c) Explain why, if $f_{xx}(a, b) > 0$ and $f_{yy}(a, b) > 0$, then $f(a, b)$ is a local minimum value for f .

When $f_{xx}(a, b)f_{yy}(a, b) - f_{xy}(a, b)^2 > 0$ and either $f_{xx}(a, b)$ or $f_{yy}(a, b)$ is negative, a slight modification of the preceding argument leads to the fact that f has a local maximum at (a, b) (the details are left to the reader). Therefore, we have proved the Second Derivative Test for functions of two variables!

Project Activity 31.5. Use the Hessian to classify the local maxima, minima, and saddle points of $f(x, y) = x^4 + y^4 - 4xy + 1$. Draw a graph of f to illustrate.

Section 32

Quadratic Forms and the Principal Axis Theorem

Focus Questions

By the end of this section, you should be able to give precise and thorough answers to the questions listed below. You may want to keep these questions in mind to focus your thoughts as you complete the section.

- What is a quadratic form on \mathbb{R}^n ?
- What does the Principal Axis Theorem tell us about quadratic forms?

Application: The Tennis Racket Effect

Try an experiment with a tennis racket (or a squash racket, or a ping pong paddle). Let us define a 3D coordinate system with the center of the racket as the origin with the head of the racket lying in the xy -plane. We let \mathbf{u}_1 be the vector in the direction of the handle and \mathbf{u}_2 the perpendicular direction (still lying in the plane defined by the head) as illustrated in Figure 32.1. We then let \mathbf{u}_3 be a vector perpendicular to the plane of the head. Hold the racket by the handle and spin it to make one rotation around the \mathbf{u}_1 axis. This is pretty easy. It is also not difficult to throw the racket so that it rotates around the \mathbf{u}_3 . Now toss the racket into the air to make one complete rotation around the axis of the vector \mathbf{u}_2 and catch the handle. Repeat this several times. You should notice that in most instances, the racket will also have made a half rotation around the \mathbf{u}_1 axis so that the other face of the racket now points up. This is quite different than the rotations around the \mathbf{u}_1 and \mathbf{u}_3 axes. A good video that illustrates this behavior can be seen at <https://www.youtube.com/watch?v=4dqCQqI-Gis>.

This effect is a result in classical mechanics that describes the rotational movement of a rigid body in space, called the *tennis racket effect* (or the Dzhanibekov effect, after the Russian cosmonaut Vladimir Dzhanibekov who discovered the theorem's consequences while in zero gravity in space – you can see an illustration of this in the video at https://www.youtube.com/watch?v=L2o9eB1_Gzw). The result is simple to see in practice, but is difficult to intuitively

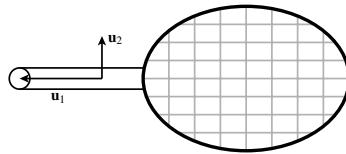


Figure 32.1: Two principal axes of a tennis racket.

understand why the behavior is different around the intermediate axis. There is a story of a student who asked the famous physicist Richard Feynman if there is any intuitive way to understand the result; Feynman supposedly went into deep thought for about 10 or 15 seconds and answered, "no." As we will see later in this section, we can understand this effect using the principal axes of a rigid body.

Introduction

We are familiar with quadratic equations in algebra. Examples of quadratic equations include $x^2 = 1$, $x^2 + y^2 = 1$, and $x^2 + xy + y^2 = 3$. We don't, however, have to restrict ourselves to two variables. A quadratic equation in n variables is any equation in which the sum of the exponents in any monomial term is 2. So a quadratic equation in the variables x_1, x_2, \dots, x_n is an equation in the form

$$\begin{aligned} & a_{11}x_1^2 + a_{22}x_2^2 + a_{33}x_3^2 + \cdots + a_{nn}x_n^2 \\ & + a_{12}x_1x_2 + a_{13}x_1x_3 + \cdots + a_{1n}x_1x_n \\ & + a_{23}x_2x_3 + a_{24}x_2x_4 + \cdots + a_{2n}x_2x_n + \cdots \\ & + a_{n-1n}x_{n-1}x_n \\ & = c \end{aligned}$$

for some constant c . In matrix notation this expression on the left of this equation has the form

$$\sum_{i=1}^n \sum_{j=1}^n a_{ij}x_i x_j = \mathbf{x}^\top A \mathbf{x}$$

where $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$ and A is the $n \times n$ matrix $A = [a_{ij}]$. For example, if $A = \begin{bmatrix} 1 & 3 & -2 \\ -1 & 1 & 2 \\ 0 & 2 & -2 \end{bmatrix}$,

then we get the quadratic expression $x_1^2 + 3x_1x_2 - 2x_1x_3 - x_2x_1 + x_2^2 + 2x_2x_3 + 2x_3x_2 - 2x_3^2$. We should note here that the terms involving $x_i x_j$ and $x_j x_i$ are repeated in our sum, but

$$a_{ij}x_i x_j + a_{ji}x_j x_i = 2 \left(\frac{a_{ij} + a_{ji}}{2} \right) x_i x_j$$

and so we could replace a_{ij} and a_{ji} both with $\left(\frac{a_{ij} + a_{ji}}{2} \right)$ without changing the quadratic form. With this alteration in mind, we can then assume that A is a symmetric matrix. So in the previous

example, the symmetric matrix $A' = \begin{bmatrix} 1 & 1 & -1 \\ 1 & 1 & 2 \\ -1 & 2 & -2 \end{bmatrix}$ gives the same quadratic expression. This leads to the following definition.

Definition 32.1. A **quadratic form** on \mathbb{R}^n is a function Q defined by

$$Q(\mathbf{x}) = \mathbf{x}^T A \mathbf{x}$$

for some $n \times n$ symmetric matrix A .

As we show in Exercise 7., the symmetric matrix A is unique to the quadratic form, so we call the symmetric matrix A is the *matrix of the quadratic form*. It is these quadratic forms that we will study in this section.

Preview Activity 32.1.

- (1) To get a little more comfortable with quadratic forms, write the quadratic forms in matrix form, explicitly identifying the vector \mathbf{x} and the symmetric matrix A of the quadratic form.
 - (a) $3x_1^2 - 2x_2^2 + 4x_1x_2 + x_2x_3$
 - (b) $x_1x_4 + 4x_2x_3 - x_2^2 + 10x_1x_5$
- (2) Some quadratic forms form equations in \mathbb{R}^2 that are very familiar: $x^2 + y^2 = 1$ is an equation of a circle, $2x^2 + 3y^2 = 2$ is an equation of an ellipse, and $x^2 - y^2 = 1$ is an equation of a hyperbola. Of course, these do not represent all of the quadratic forms in \mathbb{R}^2 – some contain cross-product terms. We can recognize the equations above because they contain no cross-product terms (terms involving xy). We can more easily recognize the quadratic forms that contain cross-product terms if we can somehow rewrite the forms in a different format with no cross-product terms. We illustrate how this can be done with the quadratic form Q defined by $Q(\mathbf{x}) = x^2 - xy + y^2$.
 - (a) Write $Q(\mathbf{x})$ in the form $\mathbf{x}^T A \mathbf{x}$, where A is a 2×2 symmetric matrix.
 - (b) Since A is a symmetric matrix we can orthogonally diagonalize A . Given that the eigenvalues of A are $\frac{3}{2}$ and $\frac{1}{2}$ with corresponding eigenvectors $\begin{bmatrix} -1 \\ 1 \end{bmatrix}$ and $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$, respectively, find a matrix P that orthogonally diagonalizes A .
 - (c) Define $\mathbf{y} = \begin{bmatrix} w \\ z \end{bmatrix}$ to satisfy $\mathbf{x} = P\mathbf{y}$. Substitute for \mathbf{x} in the quadratic form $Q(\mathbf{x})$ to write the quadratic form in terms of w and z . What kind of graph does the quadratic equation $Q(\mathbf{x}) = 1$ have?

Equations Involving Quadratic Forms in \mathbb{R}^2

When we consider equations of the form $Q(\mathbf{x}) = d$, where Q is a quadratic form in \mathbb{R}^2 and d is a constant, we wind up with old friends like $x^2 + y^2 = 1$, $2x^2 + 3y^2 = 2$, or $x^2 - y^2 = 1$. As we saw in Preview Activity 32.1 these equations are relatively easy to recognize. However, when we have



cross-product terms, like in $x^2 - xy + y^2 = 1$, it is not so easy to identify the curve the equation represents. If there was a way to eliminate the cross-product term xy from this form, we might be more easily able to recognize its graph. The discussion in this section will focus on quadratic forms in \mathbb{R}^2 , but we will see later that the arguments work in any number of dimensions. While working in \mathbb{R}^2 we will use the standard variables x and y instead of x_1 and x_2 .

In general, the equation of the form $Q(\mathbf{x}) = d$, where Q is a quadratic form in \mathbb{R}^2 defined by a matrix $A = \begin{bmatrix} a & b/2 \\ b/2 & c \end{bmatrix}$ and d is a constant looks like

$$ax^2 + bxy + cy^2 = d.$$

The graph of an equation like this is either an ellipse (a circle is a special case of an ellipse), a hyperbola, two non-intersecting lines, a point, or the empty set (see Exercise 5.). The quadratic forms do not involve linear terms, so we don't consider the cases of parabolas. One way to see into which category one of these quadratic form equations falls is to write the equation in standard form.

The standard forms for quadratic equations in \mathbb{R}^2 are as follows, where a and b are nonzero constants and h and k are any constants.

Lines: $ax^2 = 1$ or $ay^2 = 1$ ($a > 0$)

Ellipse: $\frac{(x-h)^2}{a^2} + \frac{(y-k)^2}{b^2} = 1$

Hyperbola: $\frac{(x-h)^2}{a^2} - \frac{(y-k)^2}{b^2} = 1$ or $\frac{(y-k)^2}{b^2} - \frac{(x-h)^2}{a^2} = 1$

Preview Activity 32.1 contains the main tool that we need to convert a quadratic form into one of these standard forms. By this we mean that if we have a quadratic form Q in the variables x_1, x_2, \dots, x_n , we want to find variables y_1, y_2, \dots, y_n in terms of x_1, x_2, \dots, x_n so that when written in terms of the variables y_1, y_2, \dots, y_n the quadratic form Q contains no cross terms. In other words,

we want to find a vector $\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$ so that $Q(\mathbf{x}) = \mathbf{y}^\top D \mathbf{y}$, where D is a diagonal matrix. Since

every real symmetric matrix is orthogonally diagonalizable, we will always be able to find a matrix P that orthogonally diagonalizes A . The details are as follows.

Let Q be the quadratic form defined by $Q(\mathbf{x}) = \mathbf{x}^\top A \mathbf{x}$, where A is an $n \times n$ symmetric matrix. As in Preview Activity 32.1, the fact that A is symmetric means that we can find an orthogonal matrix $P = [\mathbf{p}_1 \ \mathbf{p}_2 \ \mathbf{p}_3 \ \cdots \ \mathbf{p}_n]$ whose columns are orthonormal eigenvectors of A corresponding to eigenvalues $\lambda_1, \lambda_2, \dots, \lambda_n$, respectively. Letting $\mathbf{y} = P^\top \mathbf{x}$ give us $\mathbf{x} = P\mathbf{y}$ and

$$\mathbf{x}^\top A \mathbf{x} = (P\mathbf{y})^\top A (P\mathbf{y}) = \mathbf{y}^\top (P^\top A P) \mathbf{y} = \mathbf{y}^\top D \mathbf{y},$$

where D is the diagonal matrix whose i th diagonal entry is λ_i .

Moreover, the set $\mathcal{B} = \{\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_n\}$ is an orthonormal basis for \mathbb{R}^n and so defines a coordinate system for \mathbb{R}^n . Note that if $\mathbf{y} = [y_1 \ y_2 \ \cdots \ y_n]^\top$, then

$$\mathbf{x} = P\mathbf{y} = y_1\mathbf{p}_1 + y_2\mathbf{p}_2 + \cdots + y_n\mathbf{p}_n.$$



Thus, the coordinate vector of \mathbf{x} with respect to \mathcal{B} is \mathbf{y} , or $[\mathbf{x}]_{\mathcal{B}} = \mathbf{y}$. We summarize in Theorem 32.2.

Theorem 32.2 (Principal Axis Theorem). *Let A be an $n \times n$ symmetric matrix. There is an orthogonal change of variables $\mathbf{x} = P\mathbf{y}$ so that the quadratic form Q defined by $Q(\mathbf{x}) = \mathbf{x}^T A \mathbf{x}$ is transformed into the quadratic form $\mathbf{y}^T D \mathbf{y}$ where D is a diagonal matrix.*

The columns of the orthogonal matrix P in the Principal Axis Theorem form an orthogonal basis for \mathbb{R}^n and are called the *principal axes* for the quadratic form Q . Also, the coordinate vector of \mathbf{x} with respect to this basis is \mathbf{y} .

Activity 32.1. Let Q be the quadratic form defined by $Q(\mathbf{x}) = 2x^2 + 4xy + 5y^2 = \mathbf{x}^T A \mathbf{x}$, where $\mathbf{x} = \begin{bmatrix} x \\ y \end{bmatrix}$ and $A = \begin{bmatrix} 2 & 2 \\ 2 & 5 \end{bmatrix}$.

- (a) The eigenvalues of A are $\lambda_1 = 6$ and $\lambda_2 = 1$ with corresponding eigenvectors $\mathbf{v}_1 = [1 \ 2]^T$ and $\mathbf{v}_2 = [-2 \ 1]^T$, respectively. Find an orthogonal matrix P with determinant 1 that diagonalizes A . Is P unique? Explain. Is there a matrix without determinant 1 that orthogonally diagonalizes A ? Explain.
- (b) Use the matrix P to write the quadratic form without the cross-product.
- (c) We can view P as a change of basis matrix from the coordinate system defined by $\mathbf{y} = P^T \mathbf{x}$ to the standard coordinate system. In other words, in the standard xy coordinate system, the quadratic form is written as $\mathbf{x}^T A \mathbf{x}$, but in the new coordinate system defined by \mathbf{y} the quadratic form is written as $(P\mathbf{y})^T A (P\mathbf{y})$. As a change of basis matrix, P performs a rotation. See if you can recall what we learned about rotation matrices and determine the angle of rotation P defines. Plot the graph of the quadratic equation $Q(\mathbf{x}) = 1$ in the new coordinate system and identify this angle on the graph. Interpret the result.

Classifying Quadratic Forms

If we draw graphs of equations of the type $z = Q(\mathbf{x})$, where Q is a quadratic form, we can see that a quadratic form whose matrix does not have 0 as an eigenvalue can take on all positive values (except at $\mathbf{x} = \mathbf{0}$) as shown at left in Figure 32.2, all negative values (except at $\mathbf{x} = \mathbf{0}$) as shown in the center of Figure 32.2, or both positive and negative values as depicted at right in Figure 32.2. We can see when these cases happen by analyzing the eigenvalues of the matrix that defines the quadratic form. Let A be a 2×2 symmetric matrix with eigenvalues λ_1 and λ_2 , and let P be a matrix that orthogonally diagonalizes A so that $P^T A P = D = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}$. If we let $\mathbf{y} = \begin{bmatrix} w \\ z \end{bmatrix} = P^T \mathbf{x}$, then

$$\begin{aligned} Q(\mathbf{x}) &= \mathbf{x}^T A \mathbf{x} \\ &= \mathbf{y}^T D \mathbf{y} \\ &= \lambda_1 w^2 + \lambda_2 z^2. \end{aligned}$$

Then $Q(\mathbf{x}) \geq 0$ if all of the eigenvalues of A are positive (with $Q(\mathbf{x}) > 0$ when $\mathbf{x} \neq \mathbf{0}$) and $Q(\mathbf{x}) \leq 0$ if all of the eigenvalues of A are negative (with $Q(\mathbf{x}) < 0$ when $\mathbf{x} \neq \mathbf{0}$). If one eigenvalue



of A is positive and the other negative, then $Q(\mathbf{x})$ will take on both positive and negative values. As a result, we classify symmetric matrices (and their corresponding quadratic forms) quadratic forms according to these behaviors.

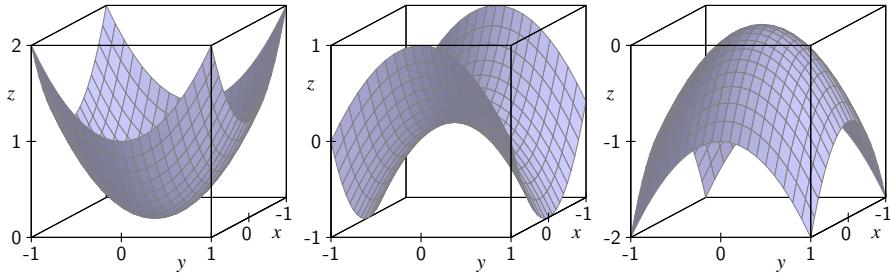


Figure 32.2: Left: Paraboloid $Q(\mathbf{x}) = x^2 + y^2$. Center: Hyperbolic Paraboloid $Q(\mathbf{x}) = x^2 - y^2$. Right: Paraboloid $Q(\mathbf{x}) = -x^2 - y^2$.

Definition 32.3. A symmetric matrix A (and its associated quadratic form Q) is

- (a) **positive definite** if $\mathbf{x}^\top A \mathbf{x} > 0$ for all $\mathbf{x} \neq \mathbf{0}$,
- (b) **positive semidefinite** if $\mathbf{x}^\top A \mathbf{x} \geq 0$ for all \mathbf{x} ,
- (c) **negative definite** if $\mathbf{x}^\top A \mathbf{x} < 0$ for all $\mathbf{x} \neq \mathbf{0}$,
- (d) **negative semidefinite** if $\mathbf{x}^\top A \mathbf{x} \leq 0$ for all \mathbf{x} ,
- (e) **indefinite** if $\mathbf{x}^\top A \mathbf{x}$ takes on both positive and negative values.

For example, the quadratic form $Q(\mathbf{x}) = x^2 + y^2$ at left in Figure 32.2 is positive definite (with repeated eigenvalue 1), the quadratic form $Q(\mathbf{x}) = -(x^2 + y^2)$ in the center of Figure 32.2 is negative definite (repeated eigenvalue -1), and the hyperbolic paraboloid $Q(\mathbf{x}) = x^2 - y^2$ at right in Figure 32.2 is indefinite (eigenvalues 1 and -1).

So we have argued that a quadratic form $Q(\mathbf{x}) = \mathbf{x}^\top A \mathbf{x}$ is positive definite if A has all positive eigenvalues, negative definite if A has all negative eigenvalues, and indefinite if A has both positive and negative eigenvalues. Similarly, the quadratic form is positive semidefinite if A has all nonnegative eigenvalues and negative semidefinite if A has all nonpositive eigenvalues. Positive definite matrices are important, as the following activity illustrates.

Activity 32.2. Let A be a symmetric $n \times n$ matrix, and define $\langle \cdot, \cdot \rangle : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$ by

$$\langle \mathbf{u}, \mathbf{v} \rangle = \mathbf{u}^\top A \mathbf{v}. \quad (32.1)$$

- (a) Explain why it is necessary for A to be positive definite in order for (32.1) to define an inner product on \mathbb{R}^n .
- (b) Show that (32.1) defines an inner product on \mathbb{R}^n if A is positive definite.

(c) Let $\langle \cdot, \cdot \rangle$ be the mapping from $\mathbb{R}^2 \times \mathbb{R}^2 \rightarrow \mathbb{R}$ defined by

$$\left\langle \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \right\rangle = 2x_1y_1 - x_1y_2 - x_2y_1 + x_2y_2.$$

Find a matrix A so that $\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{x}^\top A \mathbf{y}$ and explain why $\langle \cdot, \cdot \rangle$ defines an inner product.

Examples

What follows are worked examples that use the concepts from this section.

Example 32.4. Write the given quadratic equation in a system in which it has no cross-product terms.

(a) $8x^2 - 4xy + 5y^2 = 1$

(b) $x^2 + 4xy + y^2 = 1$

(c) $4x^2 + 4y^2 + 4z^2 + 4xy + 4xz + 4yz - 3 = 0$

Example Solution.

(a) We write the quadratic form $Q(x, y) = 8x^2 - 4xy + 5y^2$ as $\mathbf{x}^\top A \mathbf{x}$, where $\mathbf{x} = \begin{bmatrix} x \\ y \end{bmatrix}$ and $A = \begin{bmatrix} 8 & -2 \\ -2 & 5 \end{bmatrix}$. The eigenvalues for A are 9 and 4, and bases for the corresponding eigenspaces E_9 and E_4 are $\{-2, 1\}^\top$ and $\{1, 2\}^\top$, respectively. An orthogonal matrix P that orthogonally diagonalizes A is

$$P = \begin{bmatrix} -\frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \end{bmatrix}.$$

If $\mathbf{y} = [u \ v]^\top$ and we let $\mathbf{x} = P\mathbf{y}$, then we can rewrite the quadratic equation $8x^2 - 4xy + 5y^2 = 1$ as

$$\begin{aligned} 8x^2 - 4xy + 5y^2 &= 1 \\ \mathbf{x}^\top A \mathbf{x} &= 1 \\ (P\mathbf{y})^\top A (P\mathbf{y}) &= 1 \\ \mathbf{y}^\top P^\top A P \mathbf{y} &= 1 \\ \mathbf{y}^\top \begin{bmatrix} 9 & 0 \\ 0 & 4 \end{bmatrix} \mathbf{y} &= 1 \\ 9u^2 + 4v^2 &= 1. \end{aligned}$$

So the quadratic equation $8x^2 - 4xy + 5y^2 = 1$ is an ellipse.



- (b) We write the quadratic form $Q(x, y) = x^2 + 4xy + y^2$ as $\mathbf{x}^\top A \mathbf{x}$, where $\mathbf{x} = \begin{bmatrix} x \\ y \end{bmatrix}$ and $A = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}$. The eigenvalues for A are 3 and -1 , and bases for the corresponding eigenspaces E_3 and E_{-1} are $\{[1 1]^\top\}$ and $\{[-1 1]^\top\}$, respectively. An orthogonal matrix P that orthogonally diagonalizes A is

$$P = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}.$$

If $\mathbf{y} = [u v]^\top$ and we let $\mathbf{x} = P\mathbf{y}$, then we can rewrite the quadratic equation $x^2 + 4xy + y^2 = 1$ as

$$\begin{aligned} x^2 + 4xy + y^2 &= 1 \\ \mathbf{x}^\top A \mathbf{x} &= 1 \\ (P\mathbf{y})^\top A (P\mathbf{y}) &= 1 \\ \mathbf{y}^\top P^\top A P \mathbf{y} &= 1 \\ \mathbf{y}^\top \begin{bmatrix} 3 & 0 \\ 0 & -1 \end{bmatrix} \mathbf{y} &= 1 \\ 3u^2 - v^2 &= 1. \end{aligned}$$

So the quadratic equation $x^2 + 4xy + y^2 = 1$ is a hyperbola.

- (c) We write the quadratic form $Q(x, y, z) = 4x^2 + 4y^2 + 4z^2 + 4xy + 4xz + 4yz$ as $\mathbf{x}^\top A \mathbf{x}$, where $\mathbf{x} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$ and $A = \begin{bmatrix} 4 & 2 & 2 \\ 2 & 4 & 2 \\ 2 & 2 & 4 \end{bmatrix}$. The eigenvalues for A are 2 and 8, and bases for the corresponding eigenspaces E_2 and E_8 are $\{[-1 0 1]^\top, [-1 1 0]^\top\}$ and $\{[1 1 1]^\top\}$, respectively. Applying the Gram-Schmidt process to the basis for E_2 gives us an orthogonal basis $\{\mathbf{w}_1, \mathbf{w}_2\}$ of E_2 , where $\mathbf{w}_1 = [-1 0 1]^\top$ and

$$\begin{aligned} \mathbf{w}_2 &= [-1 1]^\top - \frac{[-1 0 1]^\top \cdot [-1 1 0]^\top}{[-1 0 1]^\top \cdot [-1 0 1]^\top} [-1 0 1]^\top \\ &= [-1 1 0]^\top - \frac{1}{2} [-1 0 1]^\top \\ &= \frac{1}{2} [-1 2 - 1]^\top. \end{aligned}$$

An orthogonal matrix P that orthogonally diagonalizes A is

$$P = \begin{bmatrix} -\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ 0 & \frac{2}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \end{bmatrix}.$$

If $\mathbf{y} = [u v w]^\top$ and we let $\mathbf{x} = P\mathbf{y}$, then we can rewrite the quadratic equation $4x^2 +$

$4y^2 + 4z^2 + 4xy + 4xz + 4yz = 3$ as

$$\begin{aligned} 4x^2 + 4y^2 + 4z^2 + 4xy + 4xz + 4yz &= 3 \\ \mathbf{x}^\top A \mathbf{x} &= 3 \\ (P\mathbf{y})^\top A (P\mathbf{y}) &= 3 \\ \mathbf{y}^\top P^\top A P \mathbf{y} &= 3 \\ \mathbf{y}^\top \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 8 \end{bmatrix} \mathbf{y} &= 3 \\ 2u^2 + 2v^2 + 8w^2 &= 3. \end{aligned}$$

So the quadratic equation $4x^2 + 4y^2 + 4z^2 + 4xy + 4xz + 4yz - 3 = 0$ is a ellipsoid.

Example 32.5. Let A and B be positive definite matrices, and let $C = \begin{bmatrix} 5 & -3 \\ -3 & 3 \end{bmatrix}$.

- (a) Must A be invertible? Justify your answer.
- (b) Must A^{-1} be positive definite? Justify your answer.
- (c) Must A^2 be positive definite? Justify your answer.
- (d) Must $A + B$ be positive definite? Justify your answer.
- (e) Is C positive definite? Justify your answer.

Example Solution.

- (a) Since A has all positive eigenvalues and $\det(A)$ is the product of the eigenvalues of A , then $\det(A) > 0$. Thus, A is invertible.
- (b) The fact that A is positive definite means that A is also symmetric. Recall that $(A^{-1})^\top = (A^\top)^{-1}$. Since A is symmetric, it follows that $(A^{-1})^\top = A^{-1}$ and A^{-1} is symmetric. The eigenvalues of A^{-1} are the reciprocals of the eigenvalues of A . Since the eigenvalues of A are all positive, so are the eigenvalues of A^{-1} . Thus, A^{-1} is positive definite.

- (c) Notice that

$$(A^2)^\top = (AA)^\top = A^\top A^\top = A^2,$$

so A^2 is symmetric. The eigenvalues of A^2 are the squares of the eigenvalues of A . Since no eigenvalue of A is 0, the eigenvalues of A^2 are all positive and A^2 is positive definite.

- (d) We know that B is symmetric, and

$$(A + B)^\top = A^\top + B^\top = A + B,$$

so $A + B$ is symmetric. Also, the fact that $\mathbf{x}^\top A \mathbf{x} > 0$ and $\mathbf{x}^\top B \mathbf{x} > 0$ for all \mathbf{x} implies that

$$\mathbf{x}^\top (A + B) \mathbf{x} = \mathbf{x}^\top A \mathbf{x} + \mathbf{x}^\top B \mathbf{x} > 0$$

for all \mathbf{x} . Thus, $A + B$ is positive definite.



- (e) The matrix C is symmetric and

$$\det(C - \lambda I_2) = (5 - \lambda)(3 - \lambda) - 9 = \lambda^2 - 8\lambda + 6.$$

So the eigenvalues of C are $4 + \sqrt{10}$ and $4 - \sqrt{10} \approx 0.8$. Since the eigenvalues of C are both positive, C is positive definite.

Summary

- A quadratic form on \mathbb{R}^n is a function Q defined by

$$Q(\mathbf{x}) = \mathbf{x}^\top A \mathbf{x}$$

for some $n \times n$ symmetric matrix A .

- The Principal Axis Theorem tells us that there is a change of variable $\mathbf{x} = P\mathbf{y}$ that will remove the cross-product terms from a quadratic form and allow us to identify the form and determine the principal axes for the form.

Exercises

- (1) Find the matrix for each quadratic form.

- (a) $x_1^2 - 2x_1x_2 + 4x_2^2$ if \mathbf{x} is in \mathbb{R}^2
- (b) $10x_1^2 + 4x_1x_3 + 2x_2x_3 + x_3^2$ if \mathbf{x} is in \mathbb{R}^3
- (c) $2x_1x_2 + 2x_1x_3 - x_1x_4 + 5x_2^2 + 4x_3x_4 + 8x_4^2$ if \mathbf{x} is in \mathbb{R}^4

- (2) For each quadratic form, identify the matrix A of the form, find a matrix P that orthogonally diagonalizes A , and make a change of variable that transforms the quadratic form into one with no cross-product terms.

- (a) $x_1^2 + 2x_1x_2 + x_2^2$
- (b) $-2x_1^2 + 2x_1x_2 + 4x_1x_3 - 2x_2^2 - 4x_2x_3 - x_3^2$
- (c) $11x_1^2 - 12x_1x_2 - 12x_1x_3 - 12x_1x_4 - x_2^2 - 2x_3x_4$

- (3) One topic in multivariable calculus is constrained optimization. We can use the techniques of this section to solve certain types of constrained optimization problems involving quadratic forms. As an example, we will find the maximum and minimum values of the quadratic form defined by the matrix $\begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$ on the unit circle.

- (a) First we determine some bounds on the values of a quadratic form. Let Q be the quadratic form defined by the $n \times n$ real symmetric matrix A . Let $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$ be the eigenvalues of A , and let P be a matrix that orthogonally diagonalizes A , with $P^\top AP = D$ as the matrix with diagonal entries $\lambda_1, \lambda_2, \dots, \lambda_n$ in order. Let $\mathbf{y} = [y_1 \ y_2 \ \dots \ y_n]^\top = P^\top [x_1 \ x_2 \ \dots \ x_n]^\top$.



i. Show that

$$Q(\mathbf{x}) = \lambda_1 y_1^2 + \lambda_2 y_2^2 + \cdots + \lambda_n y_n^2.$$

ii. Use the fact that $\lambda_1 \geq \lambda_i$ for each i and the fact that P (and P^\top) is an orthogonal matrix to show that

$$Q(\mathbf{x}) \leq \lambda_1 \|\mathbf{x}\|.$$

iii. Now show that $Q(\mathbf{x}) \geq \lambda_n \|\mathbf{x}\|$.

- (b) Use the result of part (a) to find the maximum and minimum values of the quadratic form defined by the matrix $\begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$ on the unit circle.

- (4) In this exercise we characterize the symmetric, positive definite, 2×2 matrices with real entries in terms of the entries of the matrices. Let $A = \begin{bmatrix} a & b/2 \\ b/2 & c \end{bmatrix}$ for some real numbers a, b , and c .

- (a) Assume that A is positive definite.

i. Show that a must be positive.

ii. Use the fact that the eigenvalues of A must be positive to show that $ac > b^2$. We conclude that if A is positive definite, then $a > 0$ and $ac > b^2$.

- (b) Now show that if $a > 0$ and $ac > b^2$, then A is positive definite. This will complete our classification of positive definite 2×2 matrices.

- (5) In this exercise we determine the form of

$$\mathbf{x}^\top A \mathbf{x} = 1, \quad (32.2)$$

where A is a symmetric 2×2 matrix. Let P be a matrix that orthogonally diagonalizes A and let $\mathbf{y} = P^\top \mathbf{x}$.

- (a) Substitute \mathbf{y} for $P^\top \mathbf{x}$ in the equation $\mathbf{x}^\top A \mathbf{x} = 1$. What form does the resulting equation have (write this form in terms of the eigenvalues of A)?

- (b) What kind of graph does the equation (32.2) have if A is positive definite? Why?

- (c) What kind of graph does the equation (32.2) have if A has both positive and negative eigenvalues? Why?

- (d) What kind of graph does the equation (32.2) have if one eigenvalue of A is zero and the other non-zero? Why?

- (6) Let $A = [a_{ij}]$ be a symmetric $n \times n$ matrix.

- (a) Show that $\mathbf{e}_i^\top A \mathbf{e}_j = a_{ij}$, where \mathbf{e}_i is the i th standard unit vector for \mathbb{R}^n . (This result will be useful in Exercise 7.)

- (b) Let \mathbf{u} be a unit eigenvector of A with eigenvalue λ . Find $\mathbf{u}^\top A \mathbf{u}$ in terms of λ .

- (7) Suppose A and B are symmetric $n \times n$ matrices, and let $Q_A(\mathbf{x}) = \mathbf{x}^\top A \mathbf{x}$ and $Q_B(\mathbf{x}) = \mathbf{x}^\top B \mathbf{x}$. If $Q_A(\mathbf{x}) = Q_B(\mathbf{x})$ for all \mathbf{x} in \mathbb{R}^n , show that $A = B$. (Hint: Use Exercise 6 (a) to compare $Q_A(\mathbf{e}_i)$ and $Q_B(\mathbf{e}_i)$, then compare $Q_A(\mathbf{e}_i + \mathbf{e}_j)$ to $Q_B(\mathbf{e}_i + \mathbf{e}_j)$ for $i \neq j$.) Thus, quadratic forms are uniquely determined by their symmetric matrices.



- (8) In this exercise we analyze all inner products on \mathbb{R}^n . Let $\langle \cdot, \cdot \rangle$ be an inner product on \mathbb{R}^n . Let $\mathbf{x} = [x_1 \ x_2 \ \dots \ x_n]^\top$ and $\mathbf{y} = [y_1 \ y_2 \ \dots \ y_n]^\top$ be arbitrary vectors in \mathbb{R}^n . Then

$$\mathbf{x} = \sum_{i=1}^n x_i \mathbf{e}_i \quad \text{and} \quad \mathbf{y} = \sum_{j=1}^n y_j \mathbf{e}_j,$$

where \mathbf{e}_i is the i th standard vector in \mathbb{R}^n .

- (a) Explain why

$$\langle \mathbf{x}, \mathbf{y} \rangle = \sum_{i=1}^n \sum_{j=1}^n x_i \langle \mathbf{e}_i, \mathbf{e}_j \rangle y_j. \quad (32.3)$$

- (b) Calculate the matrix product

$$\mathbf{x}^\top \begin{bmatrix} \langle \mathbf{e}_1, \mathbf{e}_1 \rangle & \langle \mathbf{e}_1, \mathbf{e}_2 \rangle & \cdots & \langle \mathbf{e}_1, \mathbf{e}_n \rangle \\ \langle \mathbf{e}_2, \mathbf{e}_1 \rangle & \langle \mathbf{e}_2, \mathbf{e}_2 \rangle & \cdots & \langle \mathbf{e}_2, \mathbf{e}_n \rangle \\ \vdots & \vdots & & \vdots \\ \langle \mathbf{e}_n, \mathbf{e}_1 \rangle & \langle \mathbf{e}_n, \mathbf{e}_2 \rangle & \cdots & \langle \mathbf{e}_n, \mathbf{e}_n \rangle \end{bmatrix} \mathbf{y}$$

and compare to (32.3). What do you notice?

- (c) Explain why any inner product on \mathbb{R}^n is of the form $\mathbf{x}^\top A \mathbf{y}$ for some symmetric, positive definite matrix A .
- (9) Label each of the following statements as True or False. Provide justification for your response.
- (a) **True/False** If Q is a quadratic form, then there is exactly one matrix A such that $Q(\mathbf{x}) = \mathbf{x}^\top A \mathbf{x}$.
 - (b) **True/False** The matrix of a quadratic form is unique.
 - (c) **True/False** If the matrix of a quadratic form is a diagonal matrix, then the quadratic form has no cross-product terms.
 - (d) **True/False** The eigenvectors of the symmetric matrix A form the principal axes of the quadratic form $\mathbf{x}^\top A \mathbf{x}$.
 - (e) **True/False** The principal axes of a quadratic form are orthogonal.
 - (f) **True/False** If a and c are positive, then the quadratic equation $ax^2 + bxy + cy^2 = 1$ defines an ellipse.
 - (g) **True/False** If the entries of a symmetric matrix A are all positive, then the quadratic form $\mathbf{x}^\top A \mathbf{x}$ is positive definite.
 - (h) **True/False** If a quadratic form $\mathbf{x}^\top A \mathbf{x}$ defined by a symmetric matrix A is positive definite, then the entries of A are all non-negative.
 - (i) **True/False** If a quadratic form $Q(\mathbf{x})$ on \mathbb{R}^2 is positive definite, then the graph of $z = Q(\mathbf{x})$ is a paraboloid opening upward.
 - (j) **True/False** If a quadratic form $Q(\mathbf{x})$ on \mathbb{R}^2 is negative definite, then the graph of $z = Q(\mathbf{x})$ is a paraboloid opening downward.



- (k) **True/False** If a quadratic form $Q(\mathbf{x})$ on \mathbb{R}^2 is indefinite, then there is a nonzero vector \mathbf{x} such that $Q(\mathbf{x}) = 0$.
- (l) **True/False** If $Q(\mathbf{x})$ is positive definite, then so is the quadratic form $aQ(\mathbf{x})$ for $a > 0$.
- (m) **True/False** If $Q(\mathbf{x}) = \mathbf{x}^T A \mathbf{x}$ is indefinite, then at least one of the eigenvalues of A is negative and at least one positive.
- (n) **True/False** If $n \times n$ symmetric matrices A and B define positive definite quadratic forms, then so does $A + B$.
- (o) **True/False** If an invertible symmetric matrix A defines a positive definite quadratic form, then so does A^{-1} .

Project: The Tennis Racket Theorem

If a particle of mass m and velocity v is moving in a straight line, its kinetic energy KE is given by $KE = \frac{1}{2}mv^2$. If, instead, the particle rotates around an axis with angular velocity ω (in radians per unit of time), its linear velocity is $v = r\omega$, where r is the radius of the particle's circular path. Substituting into the kinetic energy formula shows that the kinetic energy of the rotating particle is then $KE = \frac{1}{2}(mr^2)\omega^2$. The quantity mr^2 is called the *moment of inertia* of the particle and is denoted by I . So $KE = \frac{1}{2}I\omega^2$ for a rotating particle. Notice that the larger the value of r , the larger the inertia. You can imagine this with a figure skater. When a skater spins along their major axis with their arms outstretched, the speed at which they rotate is lower than when they bring their arms into their bodies. The moment of inertia for rotational motion plays a role similar to the mass in linear motion. Essentially, the inertia tells us how resistant the particle is to rotation.

To understand the tennis racket effect, we are interested in rigid bodies as they move through space. Any rigid body in three space has three principal axes about which it likes to spin. These axes are at right angles to each other and pass through the center of mass. Think of enclosing the object in an ellipsoid – the longest axis is the *primary* axis, the middle axis is the *intermediate* axis, and the third axis is the *third* axis. As a rigid body moves through space, it rotates around these axes and there is inertia along each axis. Just like with a tennis racket, if you were to imagine an axle along any of the principal axes and spin the object along that axel, it will either rotate happily with no odd behavior like flipping, or it won't. The former behavior is that of a stable axis and the latter an unstable axis. The Tennis Racket Theorem is a statement about the rotation of the body. Essentially, the Tennis Racket Theorem states that the rotation of a rigid object around its primary and third principal axes is stable, while rotation around its intermediate axis is not. To understand why this is so, we need to return to moments of inertia.

Assume that we have a rigid body moving through space. Euler's (rotation) equation describes the rotation of a rigid body with respect to the body's principal axes of inertia. Assume that I_1 , I_2 , and I_3 are the moments of inertia around the primary, intermediate, and third principal axes with $I_1 > I_2 > I_3$. Also assume that ω_1 , ω_2 , and ω_3 are the components of the angular velocity along each axis. When there is no torque applied, using a principal orthogonal coordinates, Euler's

equation tells us that

$$I_1\dot{\omega}_1 = (I_2 - I_3)\omega_2\omega_3 \quad (32.4)$$

$$I_2\dot{\omega}_2 = (I_3 - I_1)\omega_3\omega_1 \quad (32.5)$$

$$I_3\dot{\omega}_3 = (I_1 - I_2)\omega_1\omega_2. \quad (32.6)$$

(The dots indicate a derivative with respect to time, which is common notation in physics.) We will use Euler's equations to understand the Tennis Racket Theorem.

Project Activity 32.1. To start, we consider rotation around the first principal axis. Our goal is to show that rotation around this axis is stable. That is, small perturbations in angular velocity will have only small effects on the rotation of the object. So we assume that ω_2 and ω_3 are small. In general, the product of two small quantities will be much smaller, so (32.4) implies that $\dot{\omega}_1$ must be very small. So we can disregard $\dot{\omega}_1$ in our calculations.

- (a) Differentiate (32.5) with respect to time to explain why

$$I_2\ddot{\omega}_2 \approx (I_3 - I_1)\dot{\omega}_3\omega_1.$$

- (b) Substitute for $\dot{\omega}_3$ from (32.6) to show that ω_2 is an approximate solution to

$$\ddot{\omega}_2 = -k\omega_2 \quad (32.7)$$

for some positive constant k .

- (c) The equation (32.7) is a differential equation because it is an equation that involves derivatives of a function. Show by differentiating twice that, if

$$\omega_2 = A \cos(\sqrt{k}t + B) \quad (32.8)$$

(where A and B are any scalars), then ω_2 is a solution to (32.7). (In fact, ω_2 is the general solution to (32.7), which is verified in just about any course in differential equations.)

Equation 32.8 shows that ω_2 is bounded, so that any slight perturbations in angular velocity have a limited effect on ω_2 . A similar argument can be made for ω_3 . This implies that the rotation around the principal axes is stable – slight changes in angular velocity have limited effects on the rotations around the other axes.

We can make a similar argument for rotation around the third principal axes.

Project Activity 32.2. In this activity, repeat the process from Project Activity to show that rotation around the third principal axis is stable. So assume that ω_1 and ω_3 are small, which implies by (32.6) implies that $\dot{\omega}_3$ must be very small and can be disregarded in calculations.

Now the issue is why is rotation around the second principal axis different.

Project Activity 32.3. Now assume that ω_1 and ω_3 are small. Thus, $\dot{\omega}_2$ is very small by (32.5), and we consider $\dot{\omega}_2$ to be negligible.

- (a) Differentiate (32.4) to show that

$$I_1\ddot{\omega}_1 \approx (I_2 - I_3)\omega_2\dot{\omega}_3.$$



- (b) Substitute for $\dot{\omega}_3$ from (32.6) to show that ω_1 is an approximate solution to

$$\ddot{\omega}_1 = k\omega_1 \quad (32.9)$$

for some positive scalar k .

- (c) The fact that the constant multiplier in (32.9) is positive instead of negative as in (32.7) completely changes the type of solution. Show that

$$\omega_1 = Ae^{\sqrt{k}t+B} \quad (32.10)$$

(where A and B are any scalars) is a solution to (32.9) (and, in fact, is the general solution). Explain why this shows that rotation around the second principal axis is not stable.

Section 33

The Singular Value Decomposition

Focus Questions

By the end of this section, you should be able to give precise and thorough answers to the questions listed below. You may want to keep these questions in mind to focus your thoughts as you complete the section.

- What is the operator norm of a matrix and what does it tell us about the matrix?
- What is a singular value decomposition of a matrix? Why is a singular value decomposition important?
- How does a singular value decomposition relate fundamental subspaces connected to a matrix?
- What is an outer product decomposition of a matrix and how is it useful?

Application: Search Engines and Semantics

Effective search engines search for more than just words. Language is complex and search engines must deal with the fact that there are often many ways to express a given concept (this is called *synonymy*, that multiple words can have the same meaning), and that a single word can have multiple meanings (*polysemy*). As a consequence, a search on a word may provide irrelevant matches (e.g., searching for *derivative* could provide pages on mathematics or financial securities) or you might search for articles on *cats* but the paper you really want uses the word *felines*. A better search engine will not necessarily try to match terms, but instead retrieve information based on concept or intent. Latent Semantic Indexing (LSI) (or *Latent Semantic Analysis*), developed in the late 1980s, helps search engines determine concept and intent in order to provide more accurate and relevant results. LSI essentially works by providing underlying (latent) relationships between words (semantics) that search engines need to provide context and understanding (indexing). LSI provides a mapping of both words and documents into a lower dimensional “concept” space, and makes the search in this new space. The mapping is provided by the singular value decomposition.

Introduction

The singular value decomposition (SVD) of a matrix is an important and useful matrix decomposition. Unlike other matrix decompositions, *every* matrix has a singular value decomposition. The SVD is used in a variety of applications including scientific computing, digital signal processing, image compression, principal component analysis, web searching through latent semantic indexing, and seismology. Recall that the eigenvector decomposition of an $n \times n$ diagonalizable matrix M has the form $P^{-1}MP$, where the columns of the matrix P are n linearly independent eigenvectors of M and the diagonal entries of the diagonal matrix $P^{-1}MP$ are the eigenvalues of M . The singular value decomposition does something similar for any matrix of any size. One of the keys to the SVD is that the matrix $A^T A$ is symmetric for any matrix A .

The Operator Norm of a Matrix

Before we introduce the Singular Value Decomposition, let us work through some preliminaries to motivate the idea. The first is to provide an answer to the question “How ‘big’ is a matrix?” There are many ways to interpret and answer this question, but a substantial (and useful) answer should involve more than just the dimensions of the matrix. A good measure of the size of a matrix, which we will refer to as the norm of the matrix, should take into account the action of the linear transformation defined by the matrix on vectors. This then will lead to questions about how difficult or easy is it to solve a matrix equation $Ax = b$.

If we want to incorporate the action of a matrix A into a calculation of the norm of A , we might think of measuring how much A can change a vector x . This could lead us to using $\|Ax\|$ as some sort of measure of a norm of A . However, since $\|A(cx)\| = |c| \|Ax\|$ for any scalar c , scaling x by a large scalar will produce a large norm, so this is not a viable definition of a norm. We could instead measure the *relative* effect that A has on a vector x as $\frac{\|Ax\|}{\|x\|}$, since this ratio does not change when x is multiplied by a scalar. The largest of all of these ratios would provide a good sense of how much A can change vectors. Thus, we define the operator norm of a matrix A as follows.

Definition 33.1. The **operator norm**¹ of a matrix A is

$$\|A\| = \max_{\|x\| \neq 0} \left\{ \frac{\|Ax\|}{\|x\|} \right\}.$$

Due to the linearity of matrix multiplication, we can restrict ourselves to unit vectors for an equivalent definition of the operator norm of the matrix A as

$$\|A\| = \max_{\|x\|=1} \{\|Ax\|\}.$$

Preview Activity 33.1.

- (1) Determine $\|A\|$ if A is the zero matrix.

¹Technically this definition should be in terms of a supremum, but because the equivalent definition restricts the x 's to a compact subset, the sup is achieved and we can use max.



(2) Determine $\|I_n\|$, where I_n is the $n \times n$ identity matrix.

(3) Let $A = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}$. Find $\|A\|$. Justify your answer. (Hint: $x_1^2 + 4x_2^2 \leq 4(x_1^2 + x_2^2)$.)

(4) If P is an orthogonal matrix, what is $\|P\|$? Why?

The operator norm of a matrix tells us that how big the action of an $m \times n$ matrix is can be determined by its action on the unit sphere in \mathbb{R}^n (the unit sphere is the set of terminal point of unit vectors). Let us consider two examples.

Example 33.2. Let $A = \begin{bmatrix} 2 & 1 \\ 2 & 5 \end{bmatrix}$. We can draw a graph to see the action of A on the unit circle. A picture of the set $\{Ax : \|x\| = 1\}$ is shown in Figure 33.1.

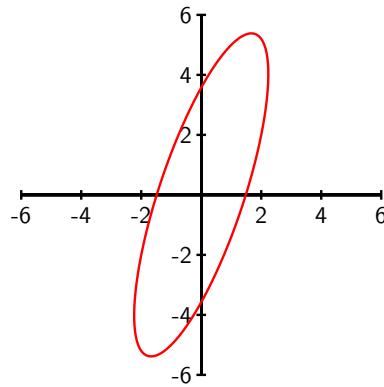


Figure 33.1: The image of the unit circle under the action of A

It appears that A transforms the unit circle into an ellipse. To find $\|A\|$ we want to maximize $\|Ax\|$ for x on the unit circle. This is the same as maximizing

$$\|Ax\|^2 = (Ax)^T(Ax) = x^T A^T A x.$$

Now $A^T A = \begin{bmatrix} 8 & 12 \\ 12 & 26 \end{bmatrix}$ is a symmetric matrix, so we can orthogonally diagonalize $A^T A$. The eigenvalues of $A^T A$ are 32 and 2. Let $P = [\mathbf{u}_1 \ \mathbf{u}_2]$, where $\mathbf{u}_1 = \begin{bmatrix} \frac{\sqrt{5}}{5} & \frac{2\sqrt{5}}{5} \end{bmatrix}^T$ is a unit eigenvector of $A^T A$ with eigenvalue 32 and $\mathbf{u}_2 = \begin{bmatrix} -\frac{2\sqrt{5}}{5} & \frac{\sqrt{5}}{5} \end{bmatrix}^T$ is a unit eigenvector of $A^T A$ with eigenvalue 2. Then P is an orthogonal matrix such that $P^T (A^T A) P = \begin{bmatrix} 32 & 0 \\ 0 & 2 \end{bmatrix} = D$. It follows that

$$x^T (A^T A) x = x^T P D P^T x = (P^T x)^T D (P^T x).$$

Now P^T is orthogonal, so $\|P^T x\| = \|x\|$ and P^T maps the unit circle to the unit circle. Moreover, if x is on the unit circle, then $y = Px$ is also on the unit circle and $P^T y = P^T Px = x$. So every point x on the unit circle corresponds to a point Px on the unit circle. Thus, the forms $x^T (A^T A) x$ and $(P^T x)^T D (P^T x)$ take on exactly the same values over all points on the unit circle. Now we

just need to find the maximum value of $(P^T \mathbf{x})^T D (P^T \mathbf{x})$. This turns out to be relatively easy since D is a diagonal matrix.

Let's simplify the notation. Let $\mathbf{y} = P^T \mathbf{x}$. Then our job is to maximize $\mathbf{y}^T D \mathbf{y}$. If $\mathbf{y} = [y_1 \ y_2]^T$, then

$$\mathbf{y}^T D \mathbf{y} = 32y_1^2 + 2y_2^2.$$

We want to find the maximum value of this expression for \mathbf{y} on the unit circle. Note that $2y_2^2 \leq 32y_1^2$ and so

$$32y_1^2 + 2y_2^2 \leq 32y_1^2 + 32y_1^2 = 32(y_1^2 + y_2^2) = 32\|\mathbf{y}\|^2 = 32.$$

Since $[1 \ 0]^T$ is on the unit circle, the expression $32y_1^2 + 2y_2^2$ attains the value 32 at some point on the unit circle, so 32 is the maximum value of $\mathbf{y}^T D \mathbf{y}$ over all \mathbf{y} on the unit circle. While we are at it, we can similarly find the minimum value of $\mathbf{y}^T D \mathbf{y}$ for \mathbf{y} on the unit circle. Since $2y_1^2 \leq 32y_1^2$ we see that

$$32y_1^2 + 2y_2^2 \geq 2y_1^2 + 2y_2^2 = 2(y_1^2 + y_2^2) = 2\|\mathbf{y}\|^2 = 2.$$

Since the expression $\mathbf{y}^T D \mathbf{y}$ attains the value 2 at $[0 \ 1]^T$ on the unit circle, we can see that $\mathbf{y}^T D \mathbf{y}$ attains the minimum value of 2 on the unit circle.

Now we can return to the expression $\mathbf{x}^T (A^T A) \mathbf{x}$. Since $\mathbf{y}^T D \mathbf{y}$ assumes the same values as $\mathbf{x}^T (A^T A) \mathbf{x}$, we can say that the maximum value of $\mathbf{x}^T (A^T A) \mathbf{x}$ for \mathbf{x} on the unit circle is 32 (and the minimum value is 2). Moreover, the quadratic form $(P^T \mathbf{x})^T D (P^T \mathbf{x})$ assumes its maximum value when $P^T \mathbf{x} = [1 \ 0]^T$ or $[-1 \ 0]^T$. Thus, the form $\mathbf{x}^T (A^T A) \mathbf{x}$ assumes its maximum value at the vector $\mathbf{x} = P[1 \ 0]^T = \mathbf{u}_1$ or $-\mathbf{u}_1$. Similarly, the quadratic form $\mathbf{x}^T (A^T A) \mathbf{x}$ attains its minimum value at $P[0 \ 1]^T = \mathbf{u}_2$ or $-\mathbf{u}_2$. We conclude that $\|A\| = \sqrt{32}$.

Figure 33.2 shows the image of the unit circle under the action of A and the images of $A\mathbf{u}_1$ and $A\mathbf{u}_2$ where $\mathbf{u}_1, \mathbf{u}_2$ are the two unit eigenvectors of $A^T A$. The image also supports that $A\mathbf{x}$ assumes its maximum and minimum values for points on the unit circle at \mathbf{u}_1 and \mathbf{u}_2 .

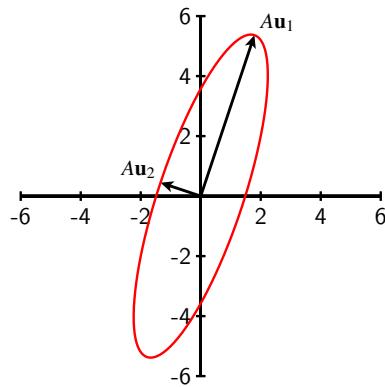


Figure 33.2: The image of the unit circle under the action of A , and the vectors $A\mathbf{u}_1$ and $A\mathbf{u}_2$

IMPORTANT NOTE 1: What we have just argued is that the maximum value of $\|A\mathbf{x}\|$ for \mathbf{x} on the unit sphere in \mathbb{R}^n is the square root of the largest eigenvalue of $A^T A$ and occurs at a corresponding unit eigenvector.

Example 33.3. This same process works for matrices other than 2×2 ones. For example, consider $A = \begin{bmatrix} -2 & 8 & 20 \\ 14 & 19 & 10 \end{bmatrix}$. In this case A maps \mathbb{R}^3 to \mathbb{R}^2 . The image of the unit sphere $\{\mathbf{x} \in \mathbb{R}^3 : \|\mathbf{x}\| = 1\}$ under left multiplication by A is a filled ellipse as shown in Figure 33.3. As with the

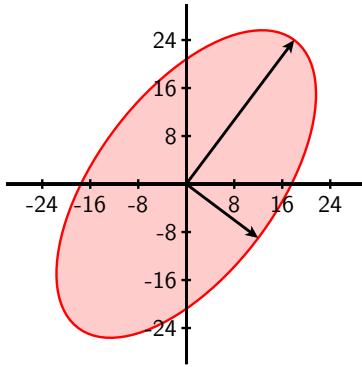


Figure 33.3: The image of the unit circle under the action of A , and the vectors $A\mathbf{u}_1$ and $A\mathbf{u}_2$

previous example, the norm of A is the square root of the maximum value of $\mathbf{x}^\top (A^\top A) \mathbf{x}$ and this

maximum value is the dominant eigenvalue of $A^\top A = \begin{bmatrix} 200 & 250 & 100 \\ 250 & 425 & 350 \\ 100 & 350 & 500 \end{bmatrix}$. The eigenvalues of

A are $\lambda_1 = 900$, $\lambda_2 = 225$, and $\lambda_3 = 0$ with corresponding unit eigenvectors $\mathbf{u}_1 = [\frac{1}{3} \frac{2}{3} \frac{2}{3}]^\top$, $\mathbf{u}_1 = [-\frac{2}{3} -\frac{1}{3} \frac{2}{3}]^\top$, and $\mathbf{u}_3 = [\frac{2}{3} -\frac{2}{3} \frac{1}{3}]^\top$. So in this case we have $\|A\| = \sqrt{900} = 30$. The transformation defined by matrix multiplication by A from \mathbb{R}^3 to \mathbb{R}^2 has a one-dimensional kernel which is spanned by the eigenvector corresponding to λ_3 . The image of the transformation is 2-dimensional and the image of the unit circle is an ellipse where $A\mathbf{u}_1$ gives the major axis of the ellipse and $A\mathbf{u}_2$ gives the minor axis. Essentially, the square roots of the eigenvalues of $A^\top A$ tell us how A stretches the image space in each direction.

IMPORTANT NOTE 2: We have just argued that the image of the unit n -sphere under the action of an $m \times n$ matrix is an ellipsoid in \mathbb{R}^m stretched the greatest amount, $\sqrt{\lambda_1}$, in the direction of an eigenvector for the largest eigenvalue (λ_1) of $A^\top A$; the next greatest amount, $\sqrt{\lambda_2}$, in the direction of a unit vector for the second largest eigenvalue (λ_2) of $A^\top A$; and so on.

Activity 33.1. Let $A = \begin{bmatrix} 0 & 5 \\ 4 & 3 \\ -2 & 1 \end{bmatrix}$. Then $A^\top A = \begin{bmatrix} 20 & 10 \\ 10 & 35 \end{bmatrix}$. The eigenvalues of $A^\top A$ are $\lambda_1 = 40$ and $\lambda_2 = 15$ with respective eigenvectors $\mathbf{v}_1 = \begin{bmatrix} \frac{1}{2} \\ 1 \end{bmatrix}$ and $\mathbf{v}_2 = \begin{bmatrix} -2 \\ 1 \end{bmatrix}$.

- (a) Find $\|A\|$.
- (b) Find a unit vector \mathbf{x} at which $\|A\mathbf{x}\|$ assumes its maximum value.

The SVD

The Singular Value Decomposition (SVD) is essentially a concise statement of what we saw in the previous section that works for *any* matrix. We will uncover the SVD in this section.

Preview Activity 33.2. Let $A = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}$. Since A is not square, we cannot diagonalize A . However, the matrix

$$A^T A = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 1 \end{bmatrix}$$

is a symmetric matrix and can be orthogonally diagonalized. The eigenvalues of $A^T A$ are 3, 1, and 0 with corresponding eigenvectors

$$\begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}, \text{ and } \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix},$$

respectively. Use appropriate technology to do the following.

- (1) Find an orthogonal matrix $V = [\mathbf{v}_1 \ \mathbf{v}_2 \ \mathbf{v}_3]$ that orthogonally diagonalizes $A^T A$, where

$$V^T (A^T A) V = \begin{bmatrix} 3 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

- (2) For $i = 1, 2$, let $\mathbf{u}_i = \frac{A\mathbf{v}_i}{\|A\mathbf{v}_i\|}$. Find each \mathbf{u}_i . Why don't we define \mathbf{u}_3 in this way?

- (3) Let $U = [\mathbf{u}_1 \ \mathbf{u}_2]$. What kind of matrix is U ? Explain.

- (4) Calculate the matrix product $U^T A V$. What do you notice? How is this similar to the eigenvector decomposition of a matrix?

Preview Activity 33.2 contains the basic ideas behind the Singular Value Decomposition. Let A be an $m \times n$ matrix with real entries. Note that $A^T A$ is a symmetric $n \times n$ matrix and, hence, it can be orthogonally diagonalized. Let $V = [\mathbf{v}_1 \ \mathbf{v}_2 \ \mathbf{v}_3 \ \cdots \ \mathbf{v}_n]$ be an $n \times n$ orthogonal matrix whose columns form an orthonormal set of eigenvectors for $A^T A$. For each i , let $(A^T A)\mathbf{v}_i = \lambda_i \mathbf{v}_i$. We know

$$V^T (A^T A) V = \begin{bmatrix} \lambda_1 & 0 & 0 & \cdots & 0 \\ 0 & \lambda_2 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & \lambda_n \end{bmatrix}.$$

Now notice that for each i we have

$$\|A\mathbf{v}_i\|^2 = (A\mathbf{v}_i)^T (A\mathbf{v}_i) = \mathbf{v}_i^T (A^T A) \mathbf{v}_i = \mathbf{v}_i^T \lambda_i \mathbf{v}_i = \lambda_i \|\mathbf{v}_i\|^2 = \lambda_i, \quad (33.1)$$

so $\lambda_i \geq 0$. Thus, the matrix $A^T A$ has no negative eigenvalues. We can always arrange the eigenvectors and eigenvalues of $A^T A$ so that

$$\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n \geq 0.$$



Also note that

$$(A\mathbf{v}_i) \cdot (A\mathbf{v}_j) = (A\mathbf{v}_i)^T (A\mathbf{v}_j) = \mathbf{v}_i^T (A^T A) \mathbf{v}_j = \mathbf{v}_i^T \lambda_j \mathbf{v}_j = \lambda_j \mathbf{v}_i \cdot \mathbf{v}_j = 0$$

if $i \neq j$. So the set $\{A\mathbf{v}_1, A\mathbf{v}_2, \dots, A\mathbf{v}_n\}$ is an orthogonal set in \mathbb{R}^m . Each of the vectors $A\mathbf{v}_i$ is in $\text{Col } A$, and so $\{A\mathbf{v}_1, A\mathbf{v}_2, \dots, A\mathbf{v}_n\}$ is an orthogonal subset of $\text{Col } A$. It is possible that $A\mathbf{v}_i = \mathbf{0}$ for some of the \mathbf{v}_i (if $A^T A$ has 0 as an eigenvalue). Let $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_r$ be the eigenvectors corresponding to the nonzero eigenvalues. Then the set

$$\mathcal{B} = \{A\mathbf{v}_1, A\mathbf{v}_2, \dots, A\mathbf{v}_r\}$$

is a linearly independent set of nonzero orthogonal vectors in $\text{Col } A$. Now we will show that \mathcal{B} is a basis for $\text{Col } A$. Let \mathbf{y} be a vector in $\text{Col } A$. Then $\mathbf{y} = A\mathbf{x}$ for some vector \mathbf{x} in \mathbb{R}^n . Recall that the vectors $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$ form an orthonormal basis of \mathbb{R}^n , so

$$\mathbf{x} = x_1 \mathbf{v}_1 + x_2 \mathbf{v}_2 + \cdots + x_n \mathbf{v}_n$$

for some scalars x_1, x_2, \dots, x_n . Since $A\mathbf{v}_j = \mathbf{0}$ for $r+1 \leq j \leq n$ we have

$$\begin{aligned} \mathbf{y} &= A\mathbf{x} \\ &= A(x_1 \mathbf{v}_1 + x_2 \mathbf{v}_2 + \cdots + x_n \mathbf{v}_n) \\ &= x_1 A\mathbf{v}_1 + x_2 A\mathbf{v}_2 + \cdots + x_r A\mathbf{v}_r + x_{r+1} A\mathbf{v}_{r+1} + \cdots + x_n A\mathbf{v}_n \\ &= x_1 A\mathbf{v}_1 + x_2 A\mathbf{v}_2 + \cdots + x_r A\mathbf{v}_r. \end{aligned}$$

So $\text{Span } \mathcal{B} = \text{Col } A$ and \mathcal{B} is an orthogonal basis for $\text{Col } A$.

Now we are ready to find the Singular Value Decomposition of A . First we create an orthonormal basis $\{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_r\}$ for $\text{Col } A$ by normalizing the vectors $A\mathbf{v}_i$. So we let

$$\mathbf{u}_i = \frac{A\mathbf{v}_i}{\|A\mathbf{v}_i\|}$$

for i from 1 to r .

Remember from (33.1) that $\|A\mathbf{v}_i\|^2 = \lambda_i$, so if we let $\sigma_i = \sqrt{\lambda_i}$, then we have

$$\mathbf{u}_i = \frac{A\mathbf{v}_i}{\sigma_i} \text{ and } A\mathbf{v}_i = \sigma_i \mathbf{u}_i.$$

We ordered the λ_i so that $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$, so we also have

$$\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0.$$

The scalars $\sigma_1, \sigma_2, \dots, \sigma_n$ are called the *singular values* of A .

Definition 33.4. Let A be an $m \times n$ matrix. The **singular values** of A are the square roots of the eigenvalues of $A^T A$.

The vectors $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_r$ are r orthonormal vectors in \mathbb{R}^m . We can extend the set $\{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_r\}$ to an orthonormal basis $\mathcal{C} = \{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_r, \mathbf{u}_{r+1}, \mathbf{u}_{r+2}, \dots, \mathbf{u}_m\}$ of \mathbb{R}^m . Recall that $A\mathbf{v}_i = \sigma_i \mathbf{u}_i$ for $1 \leq i \leq r$ and $A\mathbf{v}_j = \mathbf{0}$ for $r+1 \leq j \leq n$, so

$$\begin{aligned} AV &= A[\mathbf{v}_1 \ \mathbf{v}_2 \ \cdots \ \mathbf{v}_n] \\ &= [A\mathbf{v}_1 \ A\mathbf{v}_2 \ \cdots \ A\mathbf{v}_n] \\ &= [\sigma_1 \mathbf{u}_1 \ \sigma_2 \mathbf{u}_2 \ \cdots \ \sigma_r \mathbf{u}_r \ \mathbf{0} \ \mathbf{0} \ \cdots \ \mathbf{0}]. \end{aligned}$$



We can write the matrix $[\sigma_1 \mathbf{v}_1 \ \sigma_2 \mathbf{v}_2 \ \cdots \ \sigma_r \mathbf{v}_r \ \mathbf{0} \ \mathbf{0} \ \cdots \ \mathbf{0}]$ in another way. Let Σ be the $m \times n$ matrix defined as

$$\Sigma = \left[\begin{array}{cccc|c} \sigma_1 & & & & 0 & 0 \\ & \sigma_2 & & & & 0 \\ & & \sigma_3 & & & \\ & & & \ddots & & \\ 0 & & & & \sigma_r & 0 \\ \hline & & & & 0 & 0 \end{array} \right].$$

Now

$$[\mathbf{u}_1 \ \mathbf{u}_2 \ \cdots \ \mathbf{u}_m] \Sigma = [\sigma_1 \mathbf{u}_1 \ \sigma_1 \mathbf{u}_2 \ \cdots \ \sigma_r \mathbf{u}_r \ \mathbf{0} \ \mathbf{0} \ \cdots \ \mathbf{0}] = AV.$$

So if $U = [\mathbf{u}_1 \ \mathbf{u}_2 \ \cdots \ \mathbf{u}_m]$, then

$$U\Sigma = AV.$$

Since V is an orthogonal matrix, we have that

$$U\Sigma V^T = AVV^T = A.$$

This is the Singular Value Decomposition of A .

Theorem 33.5 (The Singular Value Decomposition). *Let A be an $m \times n$ matrix of rank r . There exist an $m \times m$ orthogonal matrix U , an $n \times n$ orthogonal matrix V , and an $m \times n$ matrix Σ whose first r diagonal entries are the singular values $\sigma_1, \sigma_2, \dots, \sigma_r$ and whose other entries are 0, such that*

$$A = U\Sigma V^T.$$

SVD Summary: A Singular Value Decomposition of an $m \times n$ matrix A of rank r can be found as follows.

- (1) Find an orthonormal basis $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \dots, \mathbf{v}_n\}$ of eigenvectors of $A^T A$ such that $(A^T A)\mathbf{v}_i = \lambda_i \mathbf{v}_i$ for i from 1 to n with $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n \geq 0$ with the first r eigenvalues being positive. The vectors $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \dots, \mathbf{v}_n$ are the *right singular vectors* of A .

- (2) Let

$$V = [\mathbf{v}_1 \ \mathbf{v}_2 \ \mathbf{v}_3 \ \cdots \ \mathbf{v}_n].$$

Then V orthogonally diagonalizes $A^T A$.

- (3) The singular values of A are the numbers σ_i , where $\sigma_i = \sqrt{\lambda_i} > 0$ for i from 1 to r . Let Σ be the $m \times n$ matrix

$$\Sigma = \left[\begin{array}{cccc|c} \sigma_1 & & & & 0 & 0 \\ & \sigma_2 & & & & 0 \\ & & \sigma_3 & & & \\ & & & \ddots & & \\ 0 & & & & \sigma_r & 0 \\ \hline & & & & 0 & 0 \end{array} \right].$$

- (4) For i from 1 to r , let $\mathbf{u}_i = \frac{A\mathbf{v}_i}{\|A\mathbf{v}_i\|}$. Then the set $\{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_r\}$ forms an orthonormal basis of $\text{Col } A$.



(5) Extend the set $\{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_r\}$ to an orthonormal basis

$$\{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_r, \mathbf{u}_{r+1}, \mathbf{u}_{r+2}, \dots, \mathbf{u}_m\}$$

of \mathbb{R}^m . Let

$$U = [\mathbf{u}_1 \ \mathbf{u}_2 \ \cdots \ \mathbf{u}_m].$$

The vectors $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_m$ are the *left singular vectors* of A .

(6) Then $A = U\Sigma V^\top$ is a singular value decomposition of A .

Activity 33.2. Let $A = \begin{bmatrix} 0 & 5 \\ 4 & 3 \\ -2 & 1 \end{bmatrix}$. Then $A^\top A = \begin{bmatrix} 20 & 10 \\ 10 & 35 \end{bmatrix}$. The eigenvalues of $A^\top A$ are $\lambda_1 = 40$ and $\lambda_2 = 15$ with respective eigenvectors $\mathbf{w}_1 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ and $\mathbf{w}_2 = \begin{bmatrix} -2 \\ 1 \end{bmatrix}$.

- (a) Find an orthonormal basis $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \dots, \mathbf{v}_n\}$ of eigenvectors of $A^\top A$. What is n ? Find the matrix V in a SVD for A .
- (b) Find the singular values of A . What is the rank r of A ? Why?
- (c) What are the dimensions of the matrix Σ in the SVD of A ? Find Σ .
- (d) Find the vectors $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_r$. If necessary, extend this set to an orthonormal basis

$$\{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_r, \mathbf{u}_{r+1}, \mathbf{u}_{r+2}, \dots, \mathbf{u}_m\}$$

of \mathbb{R}^m .

- (e) Find the matrix U so that $A = U\Sigma V^\top$ is a SVD for A .

There is another way we can write this SVD of A . Let the $m \times n$ matrix A have a singular value decomposition $U\Sigma V^\top$, where

$$U = [\mathbf{u}_1 \ \mathbf{u}_2 \ \cdots \ \mathbf{u}_m],$$

$$\Sigma = \begin{bmatrix} \sigma_1 & & & 0 & | & 0 \\ & \sigma_2 & & & | & 0 \\ & & \sigma_3 & & | & 0 \\ & & & \ddots & | & 0 \\ 0 & & & & \sigma_r & | & 0 \\ \hline & & & 0 & & | & 0 \end{bmatrix}, \text{ and}$$

$$V = [\mathbf{v}_1 \ \mathbf{v}_2 \ \mathbf{v}_3 \ \cdots \ \mathbf{v}_n].$$



Since $A = U\Sigma V^T$ we see that

$$\begin{aligned}
 A &= [\mathbf{u}_1 \ \mathbf{u}_2 \ \mathbf{u}_3 \ \cdots \ \mathbf{u}_m] \left[\begin{array}{cccc|cc} \sigma_1 & & & & 0 & 0 \\ & \sigma_2 & & & & \\ & & \sigma_3 & & & \\ & & & \ddots & & \\ 0 & & & & \sigma_r & 0 \\ \hline & & & & 0 & 0 \end{array} \right] \begin{bmatrix} \mathbf{v}_1^T \\ \mathbf{v}_2^T \\ \mathbf{v}_3^T \\ \vdots \\ \mathbf{v}_n^T \end{bmatrix} \\
 &= [\sigma_1 \mathbf{u}_1 \ \sigma_2 \mathbf{u}_2 \ \sigma_3 \mathbf{u}_3 \ \cdots \ \sigma_r \mathbf{u}_r \ \mathbf{0} \ \cdots \ \mathbf{0}] \begin{bmatrix} \mathbf{v}_1^T \\ \mathbf{v}_2^T \\ \mathbf{v}_3^T \\ \vdots \\ \mathbf{v}_n^T \end{bmatrix} \\
 &= \sigma_1 \mathbf{u}_1 \mathbf{v}_1^T + \sigma_2 \mathbf{u}_2 \mathbf{v}_2^T + \sigma_3 \mathbf{u}_3 \mathbf{v}_3^T + \cdots + \sigma_r \mathbf{u}_r \mathbf{v}_r^T. \tag{33.2}
 \end{aligned}$$

This is called an *outer product decomposition* of A and tells us everything we learned above about the action of the matrix A as a linear transformation. Each of the products $\mathbf{u}_i \mathbf{v}_i^T$ is a rank 1 matrix (see Exercise 9.), and $\|A\mathbf{v}_1\| = \sigma_1$ is the largest value A takes on the unit n -sphere, $\|A\mathbf{v}_2\| = \sigma_2$ is the next largest dilation of the unit n -sphere, and so on. An outer product decomposition allows us to approximate A with smaller rank matrices. For example, the matrix $\sigma_1 \mathbf{u}_1 \mathbf{v}_1^T$ is the best rank 1 approximation to A , $\sigma_1 \mathbf{u}_1 \mathbf{v}_1^T + \sigma_2 \mathbf{u}_2 \mathbf{v}_2^T$ is the best rank 2 approximation, and so on. This will be very useful in applications, as we will see in the next section.

SVD and the Null, Column, and Row Spaces of a Matrix

We conclude this section with a short discussion of how a singular value decomposition relates fundamental subspaces of a matrix. We have seen that the vectors $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_r$ in an SVD for an $m \times n$ matrix A form a basis for $\text{Col } A$. Recall also that $A\mathbf{v}_j = \mathbf{0}$ for $r+1 \leq j \leq n$. Since $\dim(\text{Nul } A) + \dim(\text{Col } A) = n$, it follows that the vectors $\mathbf{v}_{r+1}, \mathbf{v}_{r+2}, \dots, \mathbf{v}_n$ form a basis for $\text{Nul } A$. As you will show in the exercises, the set $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_r\}$ is a basis for $\text{Row } A$. Thus, an SVD for a matrix A tells us about three fundamental vector spaces related to A .

Examples

What follows are worked examples that use the concepts from this section.

Example 33.6. Let $A = \begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & 2 & 1 & 0 \\ 0 & 1 & 2 & 0 \end{bmatrix}$.

- (a) Find a singular value decomposition for A . You may use technology to find eigenvalues and eigenvectors of matrices.
- (b) Use the singular value decomposition to find a basis for $\text{Col } A$, $\text{Row } A$, and $\text{Nul } A$.



Example Solution.

- (a) With A as given, we have $A^T A = \begin{bmatrix} 4 & 0 & 0 & 0 \\ 0 & 5 & 4 & 0 \\ 0 & 4 & 5 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$. Technology shows that the eigenvalues of $A^T A$ are $\lambda_1 = 9$, $\lambda_2 = 4$, $\lambda_3 = 1$, and $\lambda_4 = 0$ with corresponding orthonormal eigenvectors $\mathbf{v}_1 = \frac{1}{\sqrt{2}}[0 \ 1 \ 1 \ 0]^T$, $\mathbf{v}_2 = [1 \ 0 \ 0 \ 0]^T$, $\mathbf{v}_3 = \frac{1}{\sqrt{2}}[0 \ -1 \ 1 \ 0]^T$, and $\mathbf{v}_4 = [0 \ 0 \ 0 \ 1]^T$. This makes $V = [\mathbf{v}_1 \ \mathbf{v}_2 \ \mathbf{v}_3 \ \mathbf{v}_4]$. The singular values of A are $\sigma_1 = \sqrt{9} = 3$, $\sigma_2 = \sqrt{4} = 2$, $\sigma_3 = \sqrt{1} = 1$, and $\sigma_4 = 0$, so Σ is the 3×4 matrix with the nonzero singular values along the diagonal and zeros everywhere else. Finally, we define the vectors \mathbf{u}_i as $\mathbf{u}_i = \frac{1}{\|A\mathbf{v}_i\|}A\mathbf{v}_i$. Again, technology gives us $\mathbf{u}_1 = \frac{1}{\sqrt{2}}[0 \ 1 \ 1]^T$, $\mathbf{u}_2 = [1 \ 0 \ 0]^T$, and $\mathbf{u}_3 = \frac{1}{\sqrt{2}}[0 \ -1 \ 1]^T$. Thus, a singular value decomposition of A is $U\Sigma V^T$, where

$$U = \begin{bmatrix} 0 & 1 & 0 \\ \frac{1}{\sqrt{2}} & 0 & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \end{bmatrix},$$

$$\Sigma = \begin{bmatrix} 3 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}, \text{ and}$$

$$V = \begin{bmatrix} 0 & 1 & 0 & 0 \\ \frac{1}{\sqrt{2}} & 0 & -\frac{1}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

- (b) Recall that the right singular vectors of an $m \times n$ matrix A of rank r form an orthonormal basis $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \dots, \mathbf{v}_n\}$ of eigenvectors of $A^T A$ such that $(A^T A)\mathbf{v}_i = \lambda_i \mathbf{v}_i$ for i from 1 to n with $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n \geq 0$. These vectors are the columns of the matrix $V = [\mathbf{v}_1 \ \mathbf{v}_2 \ \dots \ \mathbf{v}_n]$ in a singular value decomposition of A .

For i from 1 to r , we let $\mathbf{u}_i = \frac{A\mathbf{v}_i}{\|A\mathbf{v}_i\|}$. Then the set $\{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_r\}$ forms an orthonormal basis of $\text{Col } A$. We extend this set $\{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_r\}$ to an orthonormal basis $\{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_r, \mathbf{u}_{r+1}\mathbf{u}_{r+2}, \dots, \mathbf{u}_m\}$ of \mathbb{R}^m .

Recall also that $A\mathbf{v}_j = \mathbf{0}$ for $r+1 \leq j \leq n$. Since $\dim(\text{Nul } A) + \dim(\text{Col } A) = n$, it follows that the vectors $\mathbf{v}_{r+1}, \mathbf{v}_{r+2}, \dots, \mathbf{v}_n$ form a basis for $\text{Nul } A$. Finally, the set $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_r\}$ is a basis for $\text{Row } A$.

So in our example, we have $m = 3$, $n = 4$, $\mathbf{v}_1 = \frac{1}{\sqrt{2}}[0 \ 1 \ 1 \ 0]^T$, $\mathbf{v}_2 = [1 \ 0 \ 0 \ 0]^T$, $\mathbf{v}_3 = \frac{1}{\sqrt{2}}[0 \ -1 \ 1 \ 0]^T$, and $\mathbf{v}_4 = [0 \ 0 \ 0 \ 1]^T$. Since the singular values of A are 3, 2, 1, and 0, it follows that $r = \text{rank}(A) = 3$. We also have $\mathbf{u}_1 = \frac{1}{\sqrt{2}}[0 \ 1 \ 1]^T$, $\mathbf{u}_2 = [1 \ 0 \ 0]^T$, and $\mathbf{u}_3 = \frac{1}{\sqrt{2}}[0 \ -1 \ 1]^T$. So

$$\left\{ \frac{1}{\sqrt{2}}[0 \ 1 \ 1 \ 0]^T, [1 \ 0 \ 0 \ 0]^T, \frac{1}{\sqrt{2}}[0 \ -1 \ 1 \ 0]^T \right\}$$



is a basis for Row A and

$$\{[0 \ 0 \ 0 \ 1]^T\}$$

is a basis for Nul A . Finally, the set

$$\left\{ \frac{1}{\sqrt{2}}[0 \ 1 \ 1]^T, [1 \ 0 \ 0]^T, \frac{1}{\sqrt{2}}[[0 \ -1 \ 1]^T \right\}$$

is a basis for Col A .

Example 33.7. Let

$$A = \begin{bmatrix} 2 & 5 & 4 \\ 6 & 3 & 0 \\ 6 & 3 & 0 \\ 2 & 5 & 4 \end{bmatrix}.$$

The eigenvalues of $A^T A$ are $\lambda_1 = 144$, $\lambda_2 = 36$, and $\lambda_3 = 0$ with corresponding eigenvectors

$$\mathbf{w}_1 = \begin{bmatrix} 2 \\ 2 \\ 1 \end{bmatrix}, \quad \mathbf{w}_2 = \begin{bmatrix} -2 \\ 1 \\ 2 \end{bmatrix}, \quad \text{and} \quad \mathbf{w}_3 = \begin{bmatrix} 1 \\ -2 \\ 2 \end{bmatrix}.$$

In addition,

$$A\mathbf{w}_1 = \begin{bmatrix} 18 \\ 18 \\ 18 \\ 18 \end{bmatrix} \quad \text{and} \quad A\mathbf{w}_2 = \begin{bmatrix} 9 \\ -9 \\ -9 \\ 9 \end{bmatrix}.$$

- (a) Find orthogonal matrices U and V , and the matrix Σ , so that $U\Sigma V^T$ is a singular value decomposition of A .
- (b) Determine the best rank 1 approximation to A .

Example Solution.

- (a) Normalizing the eigenvectors \mathbf{w}_1 , \mathbf{w}_2 , and \mathbf{w}_3 to normal eigenvectors \mathbf{v}_1 , \mathbf{v}_2 , and \mathbf{v}_3 , respectively, gives us an orthogonal matrix

$$V = \begin{bmatrix} \frac{2}{3} & -\frac{2}{3} & \frac{1}{3} \\ \frac{2}{3} & \frac{1}{3} & -\frac{2}{3} \\ \frac{1}{3} & \frac{2}{3} & \frac{2}{3} \end{bmatrix}.$$

Now $A\mathbf{v}_i = A\frac{\mathbf{w}_i}{\|\mathbf{w}_i\|} = \frac{1}{\|\mathbf{w}_i\|}A\mathbf{w}_i$, so normalizing the vectors $A\mathbf{v}_1$ and $A\mathbf{v}_2$ gives us vectors

$$\mathbf{u}_1 = \frac{1}{2} \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \quad \text{and} \quad \mathbf{u}_2 = \frac{1}{2} \begin{bmatrix} 1 \\ -1 \\ -1 \\ 1 \end{bmatrix}$$



that are the first two columns of our matrix U . Given that U is a 4×4 matrix, we need to find two other vectors orthogonal to \mathbf{u}_1 and \mathbf{u}_2 that will combine with \mathbf{u}_1 and \mathbf{u}_2 to form an orthogonal basis for \mathbb{R}^4 . Letting $\mathbf{z}_1 = [1 \ 1 \ 1 \ 1]^\top$, $\mathbf{z}_2 = [1 \ -1 \ -1 \ 1]^\top$, $\mathbf{z}_3 = [1 \ 0 \ 0 \ 0]^\top$, and $\mathbf{z}_4 = [0 \ 1 \ 0 \ 1]^\top$, a computer algebra system shows that the reduced row echelon form of the matrix $[\mathbf{z}_1 \ \mathbf{z}_2 \ \mathbf{z}_3 \ \mathbf{z}_4]$ is I_4 , so that vectors $\mathbf{z}_1, \mathbf{z}_2, \mathbf{z}_3, \mathbf{z}_4$ are linearly independent. Letting $\mathbf{w}_1 = \mathbf{z}_1$ and $\mathbf{w}_2 = \mathbf{z}_2$, the Gram-Schmidt process shows that the set $\{\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3, \mathbf{w}_4\}$ is an orthogonal basis for \mathbb{R}^4 , where

$$\begin{aligned}\mathbf{w}_3 &= [1 \ 0 \ 0 \ 0]^\top - \frac{[1 \ 0 \ 0 \ 0]^\top \cdot [1 \ 1 \ 1 \ 1]^\top}{[1 \ 1 \ 1 \ 1]^\top \cdot [1 \ 1 \ 1 \ 1]^\top} [1 \ 1 \ 1 \ 1]^\top \\ &\quad - \frac{[1 \ 0 \ 0 \ 0]^\top \cdot [1 \ -1 \ -1 \ 1]^\top}{[1 \ -1 \ -1 \ 1]^\top \cdot [1 \ -1 \ -1 \ 1]^\top} [1 \ -1 \ -1 \ 1]^\top \\ &= [1 \ 0 \ 0 \ 0]^\top - \frac{1}{4} [1 \ 1 \ 1 \ 1]^\top - \frac{1}{4} [1 \ -1 \ -1 \ 1]^\top \\ &= \frac{1}{4} [2 \ 0 \ 0 \ -2]^\top\end{aligned}$$

and (using $[1 \ 0 \ 0 \ -1]^\top$ for \mathbf{w}_3)

$$\begin{aligned}\mathbf{w}_4 &= [0 \ 1 \ 0 \ 0]^\top - \frac{[0 \ 1 \ 0 \ 0]^\top \cdot [1 \ 1 \ 1 \ 1]^\top}{[1 \ 1 \ 1 \ 1]^\top \cdot [1 \ 1 \ 1 \ 1]^\top} [1 \ 1 \ 1 \ 1]^\top \\ &\quad - \frac{[0 \ 1 \ 0 \ 0]^\top \cdot [1 \ -1 \ -1 \ 1]^\top}{[1 \ -1 \ -1 \ 1]^\top \cdot [1 \ -1 \ -1 \ 1]^\top} [1 \ -1 \ -1 \ 1]^\top \\ &\quad - \frac{[0 \ 1 \ 0 \ 0]^\top \cdot [1 \ 0 \ 0 \ -1]^\top}{[1 \ 0 \ 0 \ -1]^\top \cdot [1 \ 0 \ 0 \ -1]^\top} [1 \ 0 \ 0 \ -1]^\top \\ &= [0 \ 1 \ 0 \ 0]^\top - \frac{1}{4} [1 \ 1 \ 1 \ 1]^\top + \frac{1}{4} [1 \ -1 \ -1 \ 1]^\top - \mathbf{0} \\ &= \frac{1}{4} [0 \ 2 \ -2 \ 0]^\top.\end{aligned}$$

The set $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3, \mathbf{u}_4\}$ where $\mathbf{u}_1 = \frac{1}{2}[1 \ 1 \ 1 \ 1]^\top$, $\mathbf{u}_2 = \frac{1}{2}[1 \ -1 \ -1 \ 1]^\top$, $\mathbf{u}_3 = \frac{1}{\sqrt{2}}[1 \ 0 \ 0 \ -1]^\top$ and $\mathbf{u}_4 = \frac{1}{\sqrt{2}}[0 \ 1 \ -1 \ 0]^\top$ is an orthonormal basis for \mathbb{R}^4 and we can let

$$U = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & \frac{1}{\sqrt{2}} & 0 \\ \frac{1}{2} & -\frac{1}{2} & 0 & \frac{1}{\sqrt{2}} \\ \frac{1}{2} & -\frac{1}{2} & 0 & -\frac{1}{\sqrt{2}} \\ \frac{1}{2} & \frac{1}{2} & -\frac{1}{\sqrt{2}} & 0 \end{bmatrix}.$$

The singular values of A are $\sigma_1 = \sqrt{\lambda_1} = 12$ and $\sigma_2 = \sqrt{\lambda_2} = 6$, and so

$$\Sigma = \begin{bmatrix} 12 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$



Therefore, a singular value decomposition of A is $U\Sigma V^T$ of

$$\begin{bmatrix} \frac{1}{2} & \frac{1}{2} & \frac{1}{\sqrt{2}} & 0 \\ \frac{1}{2} & -\frac{1}{2} & 0 & \frac{1}{\sqrt{2}} \\ \frac{1}{2} & -\frac{1}{2} & 0 & -\frac{1}{\sqrt{2}} \\ \frac{1}{2} & \frac{1}{2} & -\frac{1}{\sqrt{2}} & 0 \end{bmatrix} \begin{bmatrix} 12 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \frac{2}{3} & \frac{2}{3} & \frac{1}{3} \\ -\frac{2}{3} & \frac{1}{3} & \frac{2}{3} \\ \frac{1}{3} & -\frac{2}{3} & \frac{2}{3} \end{bmatrix}.$$

(b) Determine the best rank 1 approximation to A . The outer product decomposition of A is

$$A = \sigma_1 \mathbf{u}_1 \mathbf{v}_1^T + \sigma_2 \mathbf{u}_2 \mathbf{v}_2^T.$$

So the rank one approximation to A is

$$\sigma_1 \mathbf{u}_1 \mathbf{v}_1^T = 12 \left(\frac{1}{2} \right) \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \begin{bmatrix} \frac{2}{3} & \frac{2}{3} & \frac{1}{3} \end{bmatrix} = \begin{bmatrix} 4 & 4 & 2 \\ 4 & 4 & 2 \\ 4 & 4 & 2 \\ 4 & 4 & 2 \end{bmatrix}.$$

Note that the rows in this rank one approximation are the averages of the two distinct rows in the matrix A , which makes sense considering that this is the closest rank one matrix to A .

Summary

We learned about the singular value decomposition of a matrix.

- The operator norm of an $m \times n$ matrix A is

$$\|A\| = \max_{\|\mathbf{x}\| \neq 0} \frac{\|A\mathbf{x}\|}{\|\mathbf{x}\|} = \max_{\|\mathbf{x}\|=1} \|A\mathbf{x}\|.$$

The operator norm of a matrix tells us that how big the action of an $m \times n$ matrix is can be determined by its action on the unit sphere in \mathbb{R}^n .

- A singular value decomposition of an $m \times n$ matrix is of the form $A = U\Sigma V^T$, where

- $V = [\mathbf{v}_1 \mathbf{v}_2 \mathbf{v}_3 \cdots \mathbf{v}_n]$ where $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \dots, \mathbf{v}_n\}$ is an orthonormal basis of eigenvectors of $A^T A$ such that $(A^T A)\mathbf{v}_i = \lambda_i \mathbf{v}_i$ for i from 1 to n with $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n \geq 0$,
- $U = [\mathbf{u}_1 \mathbf{u}_2 \cdots \mathbf{u}_m]$ where $\mathbf{u}_i = \frac{A\mathbf{v}_i}{\|A\mathbf{v}_i\|}$ for i from 1 to r , and this orthonormal basis of $\text{Col } A$ is extended to an orthonormal basis $\{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_r, \mathbf{u}_{r+1}, \mathbf{u}_{r+2}, \dots, \mathbf{u}_m\}$ of \mathbb{R}^m ,

$$-\Sigma = \begin{array}{cc|c} \sigma_1 & & 0 & 0 \\ & \sigma_2 & & 0 \\ & & \sigma_3 & \\ & & & \ddots \\ \hline 0 & & \sigma_r & 0 \\ & 0 & & 0 \end{array}, \text{ where } \sigma_i = \sqrt{\lambda_i} > 0 \text{ for } i \text{ from 1 to } r.$$

A singular value decomposition is important in that every matrix has a singular value decomposition, and a singular value decomposition has a variety of applications including scientific computing and digital signal processing, image compression, principal component analysis, web searching through latent semantic indexing, and seismology.

- The vectors $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_r$ in an SVD for an $m \times n$ matrix A form a basis for $\text{Col } A$ while the vectors $\mathbf{v}_{r+1}, \mathbf{v}_{r+2}, \dots, \mathbf{v}_n$ form a basis for $\text{Nul } A$. Also, the set $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_r\}$ is a basis for $\text{Row } A$.
- Let A have an SVD as in the second bullet. Decomposing A as

$$A = \sigma_1 \mathbf{u}_1 \mathbf{v}_1^\top + \sigma_2 \mathbf{u}_2 \mathbf{v}_2^\top + \sigma_3 \mathbf{u}_3 \mathbf{v}_3^\top + \cdots + \sigma_r \mathbf{u}_r \mathbf{v}_r^\top$$

is an outer product decomposition of A . An outer product decomposition allows us to approximate A with smaller rank matrices. For example, the matrix $\sigma_1 \mathbf{u}_1 \mathbf{v}_1^\top$ is the best rank 1 approximation to A , $\sigma_1 \mathbf{u}_1 \mathbf{v}_1^\top + \sigma_2 \mathbf{u}_2 \mathbf{v}_2^\top$ is the best rank 2 approximation, and so on.

Exercises

- (1) Find a singular value decomposition of the following matrices.

(a) $\begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}$	(b) $\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$
(c) $\begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}$	(d) $\begin{bmatrix} 1 & 2 \\ 2 & 1 \\ 3 & 1 \\ 1 & 3 \end{bmatrix}$
(e) $\begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & 2 & 1 & 0 \\ 0 & 1 & 2 & 0 \end{bmatrix}$	

- (2) Let A be an $m \times n$ matrix of rank r with singular value decomposition $U\Sigma V^\top$, where $U = [\mathbf{u}_1 \ \mathbf{u}_2 \ \cdots \ \mathbf{u}_m]$ and $V = [\mathbf{v}_1 \ \mathbf{v}_2 \ \cdots \ \mathbf{v}_n]$. We have seen that the set $\{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_r\}$ is a basis for $\text{Col } A$, and the vectors $\mathbf{v}_{r+1}, \mathbf{v}_{r+2}, \dots, \mathbf{v}_n$ form a basis for $\text{Nul } A$. In this exercise we examine the set $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_r\}$ and determine what this set tells us about $\text{Row } A$.

- (a) Find a singular value decomposition for A^\top . (Hint: Use the singular value decomposition $U\Sigma V^\top$ for A .)
- (b) Explain why the result of (a) shows that the set $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_r\}$ is a basis for $\text{Row } A$.

(3) Let $A = \begin{bmatrix} 1 & 1 \\ 2 & 2 \\ 3 & 3 \end{bmatrix}$.

- (a) Find the singular values of A .
- (b) Find a singular value decomposition of A .
- (c) Use a singular value decomposition to find orthonormal bases for the following:



- i. $\text{Nul } A$
- ii. $\text{Col } A$
- iii. $\text{Row } A$

(4) Let A have the singular value decomposition as in (33.2).

- (a) Show, using (33.2), that $\|A\mathbf{v}_j\| = \sigma_j$.
- (b) Explain why $\|A\| = \sigma_1$. (Hint: The set $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$ is an orthonormal basis of \mathbb{R}^n . Use this to show that $\|Ax\|^2 \leq \sigma_1^2$ for any unit vector \mathbf{x} in \mathbb{R}^n .)

(5) Show that A and A^\top have the same nonzero singular values. How are their singular value decompositions related?

(6) The vectors \mathbf{v}_i that form the columns of the matrix V in a singular value decomposition of a matrix A are eigenvectors of $A^\top A$. In this exercise we investigate the vectors \mathbf{u}_i that make up the columns of the matrix U in a singular value decomposition of a matrix A for each i between 1 and the rank of A , and their connection to the matrix AA^\top .

(a) Let $A = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & -1 \end{bmatrix}$. A singular value decomposition of A is $U\Sigma V^\top$, where

$$\begin{aligned} U &= \frac{1}{\sqrt{2}} \begin{bmatrix} -1 & 1 \\ -1 & -1 \end{bmatrix}, \\ \Sigma &= \begin{bmatrix} \sqrt{3} & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}, \\ V &= \begin{bmatrix} -\frac{1}{\sqrt{6}} & -\frac{1}{3\sqrt{6}} & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \end{bmatrix}. \end{aligned}$$

- i. Determine the rank r of $A^\top A$ and identify the vectors $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_r$.
- ii. Calculate $AA^\top \mathbf{u}_i$ for each i between 1 and r . How is $AA^\top \mathbf{u}_i$ related to \mathbf{u}_i ?
- (b) Now we examine the result of part (a) in general. Let A be an arbitrary matrix. Calculate $AA^\top \mathbf{u}_i$ for $1 \leq i \leq \text{rank}(A)$ and determine specifically how $AA^\top \mathbf{u}_i$ is related to \mathbf{u}_i . What does this tell us about the vectors \mathbf{u}_i and the matrix AA^\top ?
- (c) Now show in general that the columns of U are orthonormal eigenvectors for AA^\top . (That is, what can we say about the vectors \mathbf{u}_i if $i > \text{rank}(A)$?)
- (7) If A is a symmetric matrix with eigenvalues $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n \geq 0$, what is $\|A\|$? Justify your answer.
- (8) Let A be a $n \times n$ symmetric matrix.
 - (a) Show that if \mathbf{v} is an eigenvector of A with eigenvalue λ , then \mathbf{v} is an eigenvector for $A^\top A$. What is the corresponding eigenvalue?
 - (b) Show that if \mathbf{v} is an eigenvector of $A^\top A$ with non-negative eigenvalue λ , then $A\mathbf{v}$ is an eigenvector of $A^\top A$. What is the corresponding eigenvalue?

- (c) Suppose $U\Sigma V^T$ is a singular value decomposition of A . Explain why $V\Sigma V^T$ is also a singular value decomposition of A .
- (9) Let $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_r$ and $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_r$ be the vectors found in a singular value decomposition of a matrix A , where r is the rank of A . Show that $\mathbf{u}_i\mathbf{v}_i^T$ is a rank 1 matrix for each i . (Hint: Compare to Exercise 5. in Section 31.)
- (10) Is it possible for a matrix A to have a singular value decomposition $U\Sigma V^T$ in which $U = V$? If no, explain why. If yes, determine for which matrices we can have $U = V$.
- (11) Label each of the following statements as True or False. Provide justification for your response.
- True/False** If σ is a singular value of a matrix A , then σ is an eigenvalue of A^TA .
 - True/False** A set of right singular vectors of a matrix A is also a set of left singular vectors of A^T .
 - True/False** The transpose of a singular value decomposition of a matrix A is a singular value decomposition for A^T .
 - True/False** Similar matrices have the same singular values.
 - True/False** If A is an $n \times n$ matrix, then the singular values of A^2 are the squares of the singular values of A .
 - True/False** The Σ matrix in an SVD of A is unique.
 - True/False** The matrices U, V in an SVD of A are unique.
 - True/False** If A is a positive definite matrix, then an orthogonal diagonalization of A is an SVD of A .

Project: Latent Semantic Indexing

As an elementary example to illustrate the idea behind Latent Semantic Indexing (LSI), consider the problem of creating a program to search a collection of documents for words, or words related to a given word. Document collections are usually very large, but we use a small example for illustrative purposes. A standard example that is given in several publications² is the following. Suppose we have nine documents c_1 through c_5 (titles of documents about human-computer interaction) and m_1 through m_4 (titles of graph theory papers) that make up our library:

- c_1 : *Human machine interface for ABC computer applications*
- c_2 : *A survey of user opinion of computer system response time*
- c_3 : *The EPS user interface management system*
- c_4 : *System and human system engineering testing of EPS*

² e.g., Deerwester, S., Dumais, S. T., Fumas, G. W., Landauer, T. K. and Harshman, R. Indexing by latent semantic analysis. *Journal of the American Society for Information Science*, 1990, 41: 391?407, and Landauer, T. and Dumais, S. A Solution to Plato's Problem: The Latent Semantic Analysis Theory of Acquisition, Induction, and Representation of Knowledge. *Psychological Review*, 1997. Vol. 1M. No. 2, 211-240.



- c_5 : Relation of *user* perceived *response time* to error measurement
- m_1 : The generation of random, binary, ordered *trees*
- m_2 : The intersection *graph* of paths in *trees*
- m_3 : *Graph minors IV: Widths of trees* and well-quasi-ordering
- m_4 : *Graph minors: A survey*

To make a searchable database, one might start by creating a list of key terms that appear in the documents (generally removing common words such as “a”, “the”, “of”, etc., called *stop words* – these words contribute little, if any, context). In our documents we identify the key words that are shown in italics. (Note that we are just selecting key words to make our example manageable, not necessarily identifying the most important words.) Using the key words we create a *term-document* matrix. The term-document matrix is the matrix in which the terms form the rows and the documents the columns. If $A = [a_{ij}]$ is the term-document matrix, then a_{ij} counts the number of times word i appears in document j . The term-document matrix A for our library is

$$\begin{array}{c|cccccccc} & c_1 & c_2 & c_3 & c_4 & c_5 & m_1 & m_2 & m_3 & m_4 \\ \text{human} & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ \text{interface} & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ \text{computer} & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \text{user} & 0 & 1 & 1 & 0 & 1 & 0 & 0 & 0 & 0 \\ \text{system} & 0 & 1 & 1 & 2 & 0 & 0 & 0 & 0 & 0 \\ \text{response} & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ \text{time} & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ \text{EPS} & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ \text{survey} & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ \text{trees} & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 \\ \text{graph} & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 \\ \text{minors} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \end{array} \quad (33.3)$$

One of our goals is to rate the pages in our library for relevance if we search for a query. For example, suppose we want to rate the pages for the query *survey*, *computer*. This query can be represented by the vector $\mathbf{q} = [0\ 0\ 1\ 0\ 0\ 0\ 0\ 0\ 1\ 0\ 0\ 0]^T$.

Project Activity 33.1. In a standard term-matching search with $m \times n$ term-document matrix A , a query vector \mathbf{q} would be matched with the terms to determine the number of matches. The matching counts the number of times each document agrees with the query.

- Explain why this matching is accomplished by the matrix-vector product $A^T \mathbf{q}$.
- Let $\mathbf{y} = [y_1\ y_2\ \dots\ y_n]^T = A^T \mathbf{q}$. Explain why $y_i = \cos(\theta_i) \|\mathbf{a}_i\| \|\mathbf{q}\|$, where \mathbf{a}_i is the i th column of A and θ_i is the angle between \mathbf{a}_i and \mathbf{q} .
- We can use the cosine calculation from part (b) to compare matches to our query – the closer the cosine is to 1, the better the match (dividing by the product of the norms is essentially converting all vectors to unit vectors for comparison purposes). This is often referred to as the cosine distance. Calculate the cosines of the θ_i for our example of the query $\mathbf{q} = [0\ 0\ 1\ 0\ 0\ 0\ 0\ 0\ 1\ 0\ 0\ 0]^T$. Order the documents from best to worst match for this query.



Though we were able to rate the documents in Project Activity 33.1 using the cosine distance, the result is less than satisfying. Documents c_3 , c_4 , and c_5 are all related to computers but do not appear at all in or results. This is a problem with language searches – we don’t want to compare just words, but we also need to compare the concepts the words represent. The fact that words can represent different things implies that a random choice of word by different authors can introduce noise into the word-concept relationship. To filter out this noise, we can apply the singular value decomposition to find a smaller set of concepts to better represent the relationships. Before we do so, we examine some useful properties of the term-document matrix.

Project Activity 33.2. Let $A = [\mathbf{a}_1 \ \mathbf{a}_2 \ \cdots \ \mathbf{a}_9]$, where $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_9$ are the columns of A .

- (a) In Project Activity 33.1 you should have seen that $b_{ij} = \mathbf{a}_i^\top \mathbf{a}_j = \mathbf{a}_i \cdot \mathbf{a}_j$. Assume for the moment that all of the entries in A are either 0 or 1. Explain why in this case the dot product $\mathbf{a}_i \cdot \mathbf{a}_j$ tells us how many terms documents i and j have in common. Also, the matrix $A^\top A$ takes dot products of the columns of A , which refer to what’s happening in each document and so is looking at document-document interactions. For these reasons, we call $A^\top A$ the document-document matrix.
- (b) Use appropriate technology to calculate the entries of the matrix $C = [c_{ij}] = AA^\top$. This matrix is the term-term matrix. Assume for the moment that all of the entries in A are either 0 or 1. Explain why if terms i and j occur together in k documents, then $c_{ij} = k$.

The nature of the term-term and document-document matrices makes it realistic to think about a SVD.

Project Activity 33.3. To see why a singular value decomposition might be useful, suppose our term-document matrix A has singular value decomposition $A = U\Sigma V^\top$. (Don’t actually calculate the SVD yet).

- (a) Show that the document-document matrix $A^\top A$ satisfies $A^\top A = (V\Sigma^\top)(V\Sigma^\top)^\top$. This means that we can compare document i and document j using the dot product of row i and column j of the matrix product $V\Sigma^\top$.
- (b) Show that the term-term matrix AA^\top satisfies $AA^\top = (U\Sigma)(U\Sigma)^\top$. Thus we can compare term i and term j using the dot product of row i and column j of the matrix product $U\Sigma$. (Exercise 6 shows that the columns of U are orthogonal eigenvectors of AA^\top .)

As we will see, the connection of the matrices U and V to documents and terms that we saw in Project Activity 33.3 will be very useful when we use the SVD of the term-document matrix to reduce dimensions to a “concept” space. We will be able to interpret the rows of the matrices U and V as providing coordinates for terms and documents in this space.

Project Activity 33.4. The singular value decomposition (SVD) allows us to produce new, improved term-document matrices. For this activity, use the term-document matrix A in (33.3).

- (a) Use appropriate technology to find a singular value decomposition of A so that $A = U\Sigma V^\top$. Print your entries to two decimal places (but keep as many as possible for computational purposes).



- (b) Each singular value tells us how important its semantic dimension is. If we remove the smaller singular values (the less important dimensions), we retain the important information but eliminate minor details and noise. We produce a new term-document matrix A_k by keeping the largest k of the singular values and discarding the rest. This gives us an approximation

$$A_k = \sigma_1 \mathbf{u}_1 \mathbf{v}_1^\top + \sigma_2 \mathbf{u}_2 \mathbf{v}_2^\top + \cdots + \sigma_k \mathbf{u}_k \mathbf{v}_k^\top$$

using the outer product decomposition, where $\sigma_1, \sigma_2, \dots, \sigma_k$ are the k largest singular values of A . Note that if A is an $m \times n$ matrix, letting $U_k = [\mathbf{u}_1 \mathbf{u}_2 \cdots \mathbf{u}_k]$ (an $m \times k$ matrix), Σ_k the $k \times k$ matrix with the first k singular values along the diagonal, and $V^\top = [\mathbf{v}_1 \mathbf{v}_2 \cdots \mathbf{v}_k]^\top$ (a $k \times n$ matrix), then we can also write $A_k = U_k \Sigma_k V_k^\top$. This is sometimes referred to as a reduced SVD. Find U_2 , Σ_2 , and V_2^\top , and find the new term-document matrix A_2 .

Once we have our term-document matrix, there are three basic comparisons to make: comparing terms, comparing documents, and comparing terms and documents. Term-document matrices are usually very large, with dimension being the number of terms. By using a reduced SVD we can create a much smaller approximation. In our example, the matrix A_k in Project Activity 33.4 reduces our problem to a k -dimensional space. Intuitively, we can think of LSI as representing terms as averages of all of the documents in which they appear and documents as averages of all of the terms they contain. Through this process, LSI attempts to combine the surface information in our library into a deeper abstraction (the “concept” space) that captures the mutual relationships between terms and documents.

We now need to understand how we can represent documents and terms in this smaller space where $A_k = U_k \Sigma_k V_k^\top$. Informally, we can consider the rows of U_k as representing the coordinates of each term in the lower dimensional concept space and the columns of V_k^\top as the coordinates of the documents, while the entries of Σ_k tell us how important each semantic dimension is. The dot product of two row vectors of A_k indicates how terms compare across documents. This product is $A_k A_k^\top$. Just as in Project Activity 33.3, we have $A_k A_k^\top = (U_k \Sigma_k) (U_k \Sigma_k)^\top$. In other words, if we consider the rows of $U_k \Sigma_k$ as coordinates for terms, then the dot products of these rows give us term to term comparisons. (Note that multiplying U by Σ just stretches the rows of U by the singular values according to the importance of the concept represented by that singular value.) Similarly, the dot product between columns of A provide a comparison of documents. This comparison is given by $A_k^\top A_k = (V_k \Sigma_k^\top) (V_k \Sigma_k)^\top$ (again by Project Activity 33.3). So we can consider the rows of $V \Sigma^\top$ as providing coordinates for documents.

Project Activity 33.5. We have seen how to compare terms to terms and documents to documents. The matrix A_k itself compares terms to documents. Show that $A_k = (U_k \Sigma_k^{1/2}) (V_k \Sigma_k^{1/2})^\top$, where $\Sigma_k^{1/2}$ is the diagonal matrix of the same size as Σ_k whose diagonal entries are the square roots of the corresponding diagonal entries in Σ_k . Thus, all useful comparisons of terms and documents can be made using the rows of the matrices U and V , scaled in some way by the singular values in Σ .

To work in this smaller concept space, it is important to be able to find appropriate comparisons to objects that appeared in the original search. For example, to complete the latent structure view of the system, we must also convert the original query to a representation within the new term-document system represented by A_k . This new representation is called a *pseudo-document*.



$$\begin{array}{ll}
 \text{human} & [-0.22 \quad -0.11] \\
 \text{interface} & [-0.20 \quad -0.07] \\
 \text{computer} & [-0.24 \quad 0.04] \\
 \text{user} & [-0.40 \quad 0.06] \\
 \text{system} & [-0.64 \quad -0.17] \\
 \text{response} & [-0.27 \quad 0.11] \\
 \text{time} & [-0.27 \quad 0.11] \\
 \text{EPS} & [-0.30 \quad -0.14] \\
 \text{survey} & [-0.21 \quad 0.27] \\
 \text{trees} & [-0.01 \quad 0.49] \\
 \text{graph} & [-0.04 \quad 0.62] \\
 \text{minors} & [-0.03 \quad 0.45]
 \end{array} \tag{33.4}$$

Terms in the reduced concept space.

$$\begin{bmatrix}
 c_1 & c_2 & c_3 & c_4 & c_5 & m_1 & m_2 & m_3 & m_4 \\
 -0.20 & -0.61 & -0.46 & -0.54 & -0.28 & -0.00 & -0.01 & -0.02 & -0.08 \\
 -0.06 & 0.17 & -0.13 & -0.23 & 0.11 & 0.19 & 0.44 & 0.62 & 0.53
 \end{bmatrix} \tag{33.5}$$

Documents in the reduced concept space.

Project Activity 33.6. For an original query \mathbf{q} , we start with its term vector \mathbf{a}_q (a vector in the coordinate system determined by the columns of A) and find a representation \mathbf{v}_q that we can use as a column of V^T in the document-document comparison matrix. If this representation was perfect, then it would take a real document in the original system given by A and produce the corresponding column of U if we used the full SVD. In other words, we would have $\mathbf{a}_q = U\Sigma\mathbf{v}_q^T$.

- (a) Use the fact that $A_k = U_k\Sigma_k V_k^T$, to show that $V_k = A_k^T U_k \Sigma_k^{-1}$. It follows that \mathbf{q} is transformed into the query $\mathbf{q}_k = \mathbf{q}^T U_k \Sigma_k^{-1}$.
- (b) In our example, using $k = 2$, the terms can now be represented as 2-dimensional vectors (the rows of U_2 , see (33.4)), or as points in the plane. More specifically, *human* is represented by the vector (to two decimal places) $[-0.22 \quad -0.11]^T$, *interface* by $[-0.20 \quad -0.07]^T$, etc. Similarly, the documents are represented by columns of V_2 (see (33.5)), so that the document c_1 is represented by $[-0.20 \quad -0.06]^T$, c_2 by $[-0.61 \quad 0.17]^T$, etc. From this perspective we can visualize these documents in the plane. Plot the documents and the query in the 2-dimensional concept space. Then calculate the cosine distances from the query to the documents in this space. Which documents now give the best three matches to the query? Compare the matches to your plot.

As we can see from Project Activity 33.6, the original query had no match at all with any documents except c_1 , c_2 , and m_4 . In the new concept space, the query now has some connection to every document,. So LSI has made semantic connections between the terms and documents that were not present in the original term-document matrix, which gives us better results for our search.

Section 34

Approximations Using the Singular Value Decomposition

Focus Questions

By the end of this section, you should be able to give precise and thorough answers to the questions listed below. You may want to keep these questions in mind to focus your thoughts as you complete the section.

- What is the condition number of a matrix and what does it tell us about the matrix?
- What is the pseudoinverse of a matrix?
- Why are pseudoinverses useful?
- How can we find a least squares solution to an equation $A\mathbf{x} = \mathbf{b}$?

Application: Global Positioning System

You are probably familiar with the Global Positioning System (GPS). The system allows anyone with the appropriate software to accurately determine their location at any time. The applications are almost endless, including getting real-time driving directions while in your car, guiding missiles, and providing distances on golf courses.

The GPS is a worldwide radio-navigation system owned by the US government and operated by the US Air Force. GPS is one of four global navigation satellite systems. At least twenty four GPS satellites orbit the Earth at an altitude of approximately 11,000 nautical miles. The satellites are placed so that at any time at least four of them can be accessed by a GPS receiver. Each satellite carries an atomic clock to relay a time stamp along with its position in space. There are five ground stations to coordinate and ensure that the system is working properly.

The system works by triangulation, but there is also error involved in the measurements that go into determining position. Later in this section we will see how the method of least squares can be

used to determine the receiver's position.

Introduction

A singular value decomposition has many applications, and in this section we discuss how a singular value decomposition can be used in image compression, to determine how sensitive a matrix can be to rounding errors in the process of row reduction, and to solve least squares problems.

Image Compression

The digital age has brought many new opportunities for the collection, analysis, and dissemination of information. Along with these opportunities come new difficulties as well. All of this digital information must be stored in some way and be retrievable in an efficient manner. A singular value decomposition of digitally stored information can be used to compress the information or clean up corrupted information. In this section we will see how a singular value decomposition can be used in image compression. While a singular value decomposition is normally used with very large matrices, we will restrict ourselves to small examples so that we can more clearly see how a singular value decomposition is applied.

Preview Activity 34.1. Let $A = \frac{1}{4} \begin{bmatrix} 67 & 29 & -31 & -73 \\ 29 & 67 & -73 & -31 \\ 31 & 73 & -67 & -29 \\ 73 & 31 & -29 & -67 \end{bmatrix}$. A singular value decomposition for A is $U\Sigma V^T$, where

$$U = [\mathbf{u}_1 \mathbf{u}_2 \mathbf{u}_3 \mathbf{u}_4] = \frac{1}{2} \begin{bmatrix} 1 & -1 & 1 & -1 \\ 1 & 1 & 1 & -1 \\ 1 & 1 & -1 & 1 \\ 1 & -1 & -1 & 1 \end{bmatrix},$$

$$\Sigma = \begin{bmatrix} 50 & 0 & 0 & 0 \\ 0 & 20 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix},$$

$$V = [\mathbf{v}_1 \mathbf{v}_2 \mathbf{v}_3 \mathbf{v}_4] = \frac{1}{2} \begin{bmatrix} 1 & 1 & -1 & 1 \\ 1 & -1 & -1 & -1 \\ -1 & 1 & -1 & -1 \\ -1 & -1 & -1 & 1 \end{bmatrix}.$$

- (1) Write the summands in the corresponding outer product decomposition of A .
- (2) The outer product decomposition of A writes A as a sum of rank 1 matrices (the summands $\sigma_i \mathbf{u}_i \mathbf{v}_i^T$). Each summand contains some information about the matrix A . Since σ_1 is the largest of the singular values, it is reasonable to expect that the summand $A_1 = \sigma_1 \mathbf{u}_1 \mathbf{v}_1^T$ contains the most information about A among all of the summands. To get a measure of how much information A_1 contains of A , we can think of A as simply a long vector in \mathbb{R}^{mn}

where we have folded the data into a rectangular array (we will see later why taking the norm as the norm of the vector in \mathbb{R}^{nm} makes sense, but for now, just use this definition). If we are interested in determining the error in approximating an image by a compressed image, it makes sense to use the standard norm in \mathbb{R}^{mn} to determine length and distance, which is really just the Frobenius norm that comes from the Frobenius inner product defined by

$$\langle U, V \rangle = \sum u_{ij}v_{ij},$$

where $U = [u_{ij}]$ and $V = [v_{ij}]$ are $m \times n$ matrices. So in this section all the norms for matrices will refer to the Frobenius norm. Rather than computing the distance between A_1 and A to measure the error, we are more interested in the relative error

$$\frac{\|A - A_1\|}{\|A\|}.$$

- (a) Calculate the relative error in approximating A by A_1 . What does this tell us about how much information A_1 contains about A ?
- (b) Let $A_2 = \sum_{k=1}^2 \sigma_k \mathbf{u}_k \mathbf{v}_k^\top$. Calculate the relative error in approximating A by A_2 . What does this tell us about how much information A_2 contains about A ?
- (c) Let $A_3 = \sum_{k=1}^3 \sigma_k \mathbf{u}_k \mathbf{v}_k^\top$. Calculate the relative error in approximating A by A_3 . What does this tell us about how much information A_3 contains about A ?
- (d) Let $A_4 = \sum_{k=1}^4 \sigma_k \mathbf{u}_k \mathbf{v}_k^\top$. Calculate the relative error in approximating A by A_4 . What does this tell us about how much information A_4 contains about A ? Why?

The first step in compressing an image is to digitize the image. (If you completed the image compression application project in Section 23, then this will seem familiar to you.) There are many ways to do this and we will consider one of the simplest ways and only work with gray-scale images, with the scale from 0 (black) to 255 (white). A digital image can be created by taking a small grid of squares (called pixels) and coloring each pixel with some shade of gray. The resolution of this grid is a measure of how many pixels are used per square inch. As an example, consider the 16 by 16 pixel picture of a flower shown in Figure 34.1.

To store this image pixel by pixel would require $16 \times 16 = 256$ units of storage space (1 for each pixel). If we let M be the matrix whose i, j th entry is the scale of the i, j th pixel, then M is the matrix

$$\left[\begin{array}{cccccccccccccccc} 240 & 240 & 240 & 240 & 130 & 130 & 240 & 130 & 130 & 240 & 240 & 240 & 240 & 240 & 240 & 240 \\ 240 & 240 & 240 & 130 & 175 & 175 & 130 & 175 & 175 & 130 & 240 & 240 & 240 & 240 & 240 & 240 & 240 \\ 240 & 240 & 130 & 130 & 175 & 175 & 130 & 175 & 175 & 130 & 240 & 240 & 240 & 240 & 240 & 240 & 240 \\ 240 & 130 & 175 & 175 & 130 & 175 & 175 & 130 & 175 & 175 & 130 & 240 & 240 & 240 & 240 & 240 & 240 \\ 240 & 240 & 130 & 175 & 175 & 130 & 175 & 175 & 130 & 175 & 175 & 130 & 240 & 240 & 240 & 240 & 240 \\ 240 & 240 & 130 & 175 & 175 & 130 & 175 & 175 & 130 & 175 & 175 & 130 & 240 & 240 & 240 & 240 & 240 \\ 255 & 240 & 240 & 130 & 175 & 175 & 130 & 175 & 175 & 130 & 240 & 240 & 225 & 240 & 240 & 240 & 240 \\ 240 & 240 & 130 & 175 & 175 & 130 & 175 & 175 & 130 & 175 & 175 & 130 & 240 & 225 & 255 & 240 & 240 \\ 240 & 240 & 130 & 175 & 175 & 130 & 175 & 175 & 130 & 175 & 175 & 130 & 240 & 225 & 255 & 255 & 240 \\ 240 & 240 & 130 & 175 & 175 & 130 & 175 & 175 & 130 & 175 & 175 & 130 & 240 & 225 & 255 & 255 & 240 \\ 240 & 240 & 240 & 240 & 240 & 240 & 240 & 240 & 240 & 240 & 240 & 240 & 240 & 240 & 240 & 240 & 240 \\ 240 & 240 & 240 & 240 & 240 & 240 & 240 & 240 & 240 & 240 & 240 & 240 & 240 & 240 & 240 & 240 & 240 \\ 240 & 240 & 240 & 240 & 240 & 240 & 240 & 240 & 240 & 240 & 240 & 240 & 240 & 240 & 240 & 240 & 240 \\ 240 & 240 & 240 & 240 & 240 & 240 & 240 & 240 & 240 & 240 & 240 & 240 & 240 & 240 & 240 & 240 & 240 \\ 50 & 240 & 240 & 240 & 240 & 240 & 240 & 240 & 240 & 240 & 240 & 240 & 240 & 240 & 240 & 240 & 240 \\ 240 & 75 & 240 & 240 & 240 & 240 & 240 & 240 & 240 & 240 & 240 & 240 & 240 & 240 & 240 & 240 & 240 \\ 240 & 240 & 75 & 240 & 240 & 240 & 240 & 240 & 240 & 240 & 240 & 240 & 240 & 240 & 240 & 240 & 240 \\ 75 & 75 & 75 & 75 & 75 & 75 & 75 & 75 & 75 & 75 & 75 & 75 & 75 & 75 & 75 & 75 \\ 75 & 75 & 75 & 75 & 75 & 75 & 75 & 75 & 75 & 75 & 75 & 75 & 75 & 75 & 75 & 75 \end{array} \right].$$

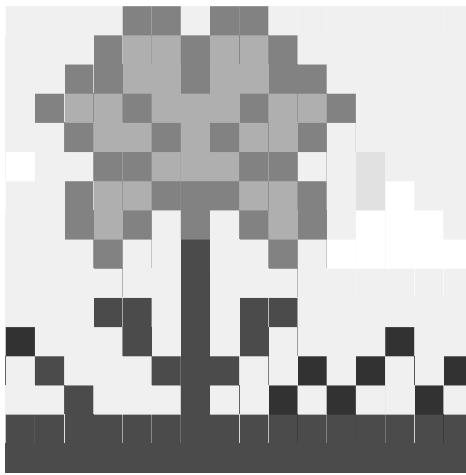


Figure 34.1: A 16 by 16 pixel image

Recall that if $U\Sigma V^T$ is a singular value decomposition for M , then we can also write M in the form

$$M = \sigma_1 \mathbf{u}_1 \mathbf{v}_1^T + \sigma_2 \mathbf{u}_2 \mathbf{v}_2^T + \sigma_3 \mathbf{u}_3 \mathbf{v}_3^T + \cdots + \sigma_{16} \mathbf{u}_{16} \mathbf{v}_{16}^T.$$

given in (33.2). For this M , the singular values are approximately

$$\left[\begin{array}{c} 3006.770088367795 \\ 439.13109000200205 \\ 382.1756550649652 \\ 312.1181752764884 \\ 254.45105800344953 \\ 203.36470770057494 \\ 152.8696215072527 \\ 101.29084240890717 \\ 63.80803769229468 \\ 39.6189181773536 \\ 17.091891798245463 \\ 12.304589419140656 \\ 4.729898943556077 \\ 2.828719409809012 \\ 6.94442317024232 \times 10^{-15} \\ 2.19689952047833 \times 10^{-15} \end{array} \right]. \quad (34.1)$$

Notice that some of these singular values are very small compared to others. As in Preview Activity 34.1, the terms with the largest singular values contain most of the information about the matrix. Thus, we shouldn't lose much information if we eliminate the small singular values. In this particular example, the last 4 singular values are significantly smaller than the rest. If we let

$$M_{12} = \sigma_1 \mathbf{u}_1 \mathbf{v}_1^T + \sigma_2 \mathbf{u}_2 \mathbf{v}_2^T + \sigma_3 \mathbf{u}_3 \mathbf{v}_3^T + \cdots + \sigma_{12} \mathbf{u}_{12} \mathbf{v}_{12}^T,$$

then we should expect the image determined by M_{12} to be close to the image made by M . The two images are presented side by side in Figure 34.2.

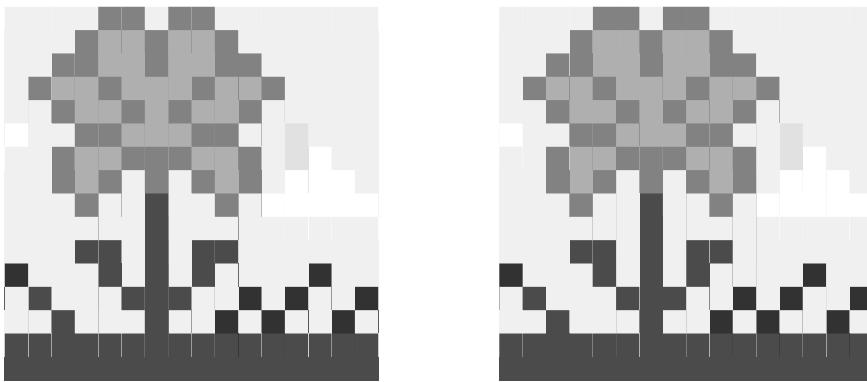


Figure 34.2: A 16 by 16 pixel image and a compressed image using a singular value decomposition.

This small example illustrates the general idea. Suppose we had a satellite image that was 1000×1000 pixels and we let M represent this image. If we have a singular value decomposition of this image M , say

$$M = \sigma_1 \mathbf{u}_1 \mathbf{v}_1^\top + \sigma_2 \mathbf{u}_2 \mathbf{v}_2^\top + \sigma_3 \mathbf{u}_3 \mathbf{v}_3^\top + \cdots + \sigma_r \mathbf{u}_r \mathbf{v}_r^\top,$$

if the rank of M is large, it is likely that many of the singular values will be very small. If we only keep s of the singular values, we can approximate M by

$$M_s = \sigma_1 \mathbf{u}_1 \mathbf{v}_1^\top + \sigma_2 \mathbf{u}_2 \mathbf{v}_2^\top + \sigma_3 \mathbf{u}_3 \mathbf{v}_3^\top + \cdots + \sigma_s \mathbf{u}_s \mathbf{v}_s^\top$$

and store the image with only the vectors $\sigma_1 \mathbf{u}_1, \sigma_2 \mathbf{u}_2, \dots, \sigma_s \mathbf{u}_s, \mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_s$. For example, if we only need 10 of the singular values of a satellite image ($s = 10$), then we can store the satellite image with only 20 vectors in \mathbb{R}^{1000} or with $20 \times 1000 = 20,000$ numbers instead of $1000 \times 1000 = 1,000,000$ numbers.

A similar process can be used to denoise data.¹

Calculating the Error in Approximating an Image

In the context where a matrix represents an image, the operator aspect of the matrix is irrelevant – we are only interested in the matrix as a holder of information. In this situation, we think of an $m \times n$ matrix as simply a long vector in \mathbb{R}^{mn} where we have folded the data into a rectangular array. If we are interested in determining the error in approximating an image by a compressed image, it makes sense to use the standard norm in \mathbb{R}^{mn} to determine length and distance. This leads to what is called the *Frobenius* norm of a matrix. The Frobenius norm $\|M\|_F$ of an $m \times n$ matrix $M = [m_{ij}]$ is

$$\|M\|_F = \sqrt{\sum m_{ij}^2}.$$

¹For example, as stated in <http://www2.imm.dtu.dk/~pch/Projekter/tsvd.html>, “The SVD [singular value decomposition] has also applications in digital signal processing, e.g., as a method for noise reduction. The central idea is to let a matrix A represent the noisy signal, compute the SVD, and then discard small singular values of A . It can be shown that the small singular values mainly represent the noise, and thus the rank- k matrix A_k represents a filtered signal with less noise.”

There is a natural corresponding inner product on the set of $m \times n$ matrices (called the *Frobenius product*) defined by

$$\langle A, B \rangle = \sum a_{ij} b_{ij},$$

where $A = [a_{ij}]$ and $B = [b_{ij}]$ are $m \times n$ matrices.² Note that

$$\|A\|_F = \sqrt{\langle A, A \rangle}.$$

If an $m \times n$ matrix M of rank r has a singular value decomposition $M = U\Sigma V^T$, we have seen that we can write M as an outer product

$$M = \sigma_1 \mathbf{u}_1 \mathbf{v}_1^T + \sigma_2 \mathbf{u}_2 \mathbf{v}_2^T + \sigma_3 \mathbf{u}_3 \mathbf{v}_3^T + \cdots + \sigma_r \mathbf{u}_r \mathbf{v}_r^T, \quad (34.2)$$

where the \mathbf{u}_i are the columns of U and the \mathbf{v}_j the columns of V . Each of the products $\mathbf{u}_i \mathbf{v}_i^T$ is an $m \times n$ matrix. Since the columns of $\mathbf{u}_i \mathbf{v}_i^T$ are all scalar multiples of \mathbf{u}_i , the matrix $\mathbf{u}_i \mathbf{v}_i^T$ is a rank 1 matrix. So (34.2) expresses M as a sum of rank 1 matrices. Moreover, if we let \mathbf{x} and \mathbf{w} be $m \times 1$ vectors and let \mathbf{y} and \mathbf{z} be $n \times 1$ vectors with $\mathbf{y} = [y_1 \ y_2 \ \dots \ y_n]^T$ and $\mathbf{z} = [z_1 \ z_2 \ \dots \ z_n]^T$, then

$$\begin{aligned} \langle \mathbf{x} \mathbf{y}^T, \mathbf{w} \mathbf{z}^T \rangle &= \langle [y_1 \mathbf{x} \ y_2 \mathbf{x} \ \dots \ y_n \mathbf{x}], [z_1 \mathbf{w} \ z_2 \mathbf{w} \ \dots \ z_n \mathbf{w}] \rangle \\ &= \sum (y_i \mathbf{x}) \cdot (z_i \mathbf{w}) \\ &= \sum (y_i z_i) (\mathbf{x} \cdot \mathbf{w}) \\ &= (\mathbf{x} \cdot \mathbf{w}) \sum (y_i z_i) \\ &= (\mathbf{x} \cdot \mathbf{w})(\mathbf{y} \cdot \mathbf{z}). \end{aligned}$$

Using the vectors from the singular value decomposition of M as in (34.2) we see that

$$\langle \mathbf{u}_i \mathbf{v}_i^T, \mathbf{u}_j \mathbf{v}_j^T \rangle = (\mathbf{u}_i \cdot \mathbf{u}_j)(\mathbf{v}_i \cdot \mathbf{v}_j) = \begin{cases} 0, & \text{if } i \neq j, \\ 1, & \text{if } i = j. \end{cases}$$

It follows that

$$\|M\|_F^2 = \sum \sigma_i^2 (\mathbf{u}_i \cdot \mathbf{u}_i)(\mathbf{v}_i \cdot \mathbf{v}_i) = \sum \sigma_i^2. \quad (34.3)$$

Activity 34.1. Verify (34.3) that $\|M\|_F^2 = \sum \sigma_i^2$.

When we used the singular value decomposition to approximate the image defined by M , we replaced M with a matrix of the form

$$M_k = \sigma_1 \mathbf{u}_1 \mathbf{v}_1^T + \sigma_2 \mathbf{u}_2 \mathbf{v}_2^T + \sigma_3 \mathbf{u}_3 \mathbf{v}_3^T + \cdots + \sigma_k \mathbf{u}_k \mathbf{v}_k^T. \quad (34.4)$$

We call M_k the rank k approximation to M . Notice that the outer product expansion in (34.4) is in fact a singular value decomposition for M_k . The error E_k in approximating M with M_k is

$$E_k = M - M_k = \sigma_{k+1} \mathbf{u}_{k+1} \mathbf{v}_{k+1}^T + \sigma_{k+2} \mathbf{u}_{k+2} \mathbf{v}_{k+2}^T + \cdots + \sigma_r \mathbf{u}_r \mathbf{v}_r^T. \quad (34.5)$$

²This is the same inner product that we defined as the Frobenius inner product in Section 29, where $\langle A, B \rangle = \text{trace}(AB^T)$.



Once again, notice that (34.5) is a singular value decomposition for E_k . We define the relative error in approximating M with M_k as

$$\frac{\|E_k\|}{\|M\|}.$$

Now (34.3) shows that

$$\frac{\|E_k\|}{\|M\|} = \sqrt{\frac{\sum_{i=k+1}^r \sigma_i^2}{\sum_{i=1}^r \sigma_i^2}}.$$

In applications, we often want to retain a certain degree of accuracy in our approximations and this error term can help us accomplish that.

In our flower example, the singular values of M are given in (34.1). The relative error in approximating M with M_{12} is

$$\sqrt{\frac{\sum_{i=13}^{16} \sigma_i^2}{\sum_{i=1}^{16} \sigma_i^2}} \approx 0.03890987666.$$

Errors (rounded to 4 decimal places) for approximating M with some of the M_k are shown in Table 34.1

k	10	9	8	7	6
$\frac{\ E_k\ }{\ M\ }$	0.0860	0.1238	0.1677	0.2200	0.2811
k	5	4	3	2	1
$\frac{\ E_k\ }{\ M\ }$	0.3461	0.4132	0.4830	0.5566	0.6307

Table 34.1: Errors in approximating M by M_k

Activity 34.2. Let M represent the flower image.

- (a) Find the relative errors in approximating M by M_{13} and M_{14} . You can use the fact that $\sum_{i=1}^{16} \sigma_i^2 \approx 4992.553293$.
- (b) About how much of the information in the image is contained in the rank 1 approximation? Explain.

The Condition Number of a Matrix

A singular value decomposition for a matrix A can tell us a lot about how difficult it is to accurately solve a system $Ax = b$. Solutions to systems of linear equations can be very sensitive to rounding as the next exercise demonstrates.

Activity 34.3. Find the solution to each of the systems.

$$(a) \begin{bmatrix} 1.0000 & 1.0000 \\ 1.0000 & 1.0005 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 2.0000 \\ 2.0050 \end{bmatrix}$$



$$(b) \begin{bmatrix} 1.000 & 1.000 \\ 1.000 & 1.001 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 2.000 \\ 2.005 \end{bmatrix}$$

Notice that a simple rounding in the (2, 2) entry of the coefficient matrix led to a significantly different solution. If there are rounding errors at any stage of the Gaussian elimination process, they can be compounded by further row operations. This is an important problem since computers can only approximate irrational numbers with rational numbers and so rounding can be critical. Finding ways of dealing with these kinds of errors is an area of on-going research in numerical linear algebra. This problem is given a name.

Definition 34.1. A matrix A is **ill-conditioned** if relatively small changes in any entries of A can produce significant changes in solutions to the system $Ax = b$.

A matrix that is not ill-conditioned is said to be *well-conditioned*. Since small changes in entries of ill-conditioned matrices can lead to large errors in computations, it is an important problem in linear algebra to have a way to measure how ill-conditioned a matrix is. This idea will ultimately lead us to the condition number of a matrix.

Suppose we want to solve the system $Ax = b$, where A is an invertible matrix. Activity 34.3 illustrates that if A is really close to being singular, then small changes in the entries of A can have significant effects on the solution to the system. So the system can be very hard to solve accurately if A is close to singular. It is important to have a sense of how “good” we can expect any calculated solution to be. Suppose we think we solve the system $Ax = b$ but, through rounding error in our calculation of A , get a solution x' so that $Ax' = b'$, where b' is not exactly b . Let Δx be the error in our calculated solution and Δb the difference between b' and b . We would like to know how large the error $\|\Delta x\|$ can be. But this isn’t exactly the right question. We could scale everything to make $\|\Delta x\|$ as large as we want. What we really need is a measure of the *relative error* $\frac{\|\Delta x\|}{\|x\|}$, or how big the error is compared to $\|x\|$ itself. More specifically, we want to know how large the relative error in Δx is compared to the relative error in Δb . In other words, we want to know how good the relative error in Δb is as a predictor of the relative error in Δx (we may have some control over the relative error in Δb , perhaps by keeping more significant digits). So we want to know if there is a best constant C such that

$$\frac{\|\Delta x\|}{\|x\|} \leq C \frac{\|\Delta b\|}{\|b\|}.$$

This best constant C is the condition number – a measure of how well the relative error in Δb predicts the relative error in Δx . How can we find C ?

Since $Ax' = b'$ we have

$$A(x + \Delta x) = b + \Delta b.$$

Distributing on the left and using the fact that $Ax = b$ gives us

$$A\Delta x = \Delta b.$$

We return for a moment to the operator norm of a matrix. This is an appropriate norm to use here since we are considering A to be a transformation. Recall that if A is an $m \times n$ matrix, we defined the operator norm of A to be

$$\|A\| = \max_{\|x\| \neq 0} \left\{ \frac{\|Ax\|}{\|x\|} \right\} = \max_{\|x\|=1} \{\|Ax\|\}.$$



One important property that the norm has is that if the product AB is defined, then

$$\|AB\| \leq \|A\| \|B\|.$$

To see why, notice that

$$\frac{\|ABx\|}{\|x\|} = \frac{\|A(Bx)\|}{\|Bx\|} \frac{\|Bx\|}{\|x\|}.$$

Now $\frac{\|A(Bx)\|}{\|Bx\|} \leq \|A\|$ and $\frac{\|Bx\|}{\|x\|} \leq \|B\|$ by the definition of the norm, so we conclude that

$$\frac{\|ABx\|}{\|x\|} \leq \|A\| \|B\|$$

for every x . Thus,

$$\|AB\| \leq \|A\| \|B\|.$$

Now we can find the condition number. From $A\Delta x = \Delta b$ we have

$$\Delta x = A^{-1}\Delta b,$$

so

$$\|\Delta x\| \leq \|A^{-1}\| \|\Delta b\|. \quad (34.6)$$

Similarly, $b = Ax$ implies that $\|b\| \leq \|A\| \|x\|$ or

$$\frac{1}{\|x\|} \leq \frac{\|A\|}{\|b\|}. \quad (34.7)$$

Combining (34.6) and (34.7) gives

$$\begin{aligned} \frac{\|\Delta x\|}{\|x\|} &\leq \frac{\|A^{-1}\| \|\Delta b\|}{\|x\|} \\ &= \|A^{-1}\| \|\Delta b\| \left(\frac{1}{\|x\|} \right) \\ &\leq \|A^{-1}\| \|\Delta b\| \frac{\|A\|}{\|b\|} \\ &= \|A^{-1}\| \|A\| \frac{\|\Delta b\|}{\|b\|}. \end{aligned}$$

This constant $\|A^{-1}\| \|A\|$ is the best bound and so is called the condition number of A .

Definition 34.2. The **condition number** of an invertible matrix A is the number $\|A^{-1}\| \|A\|$.

How does a singular value decomposition tell us about the condition number of a matrix? Recall that the maximum value of $\|Ax\|$ for x on the unit n -sphere is σ_1 . So $\|A\| = \sigma_1$. If A is an invertible matrix and $A = U\Sigma V^T$ is a singular value decomposition for A , then

$$A^{-1} = (U\Sigma V^T)^{-1} = (V^T)^{-1}\Sigma^{-1}U^{-1} = V\Sigma^{-1}U^T,$$



where

$$\Sigma^{-1} = \begin{bmatrix} \frac{1}{\sigma_1} & & & 0 \\ & \frac{1}{\sigma_2} & & \\ & & \frac{1}{\sigma_3} & \\ & & & \ddots \\ 0 & & & & \frac{1}{\sigma_n} \end{bmatrix}.$$

Now $V\Sigma^{-1}U^\top$ is a singular value decomposition for A^{-1} with the diagonal entries in reverse order, so

$$\|A^{-1}\| = \frac{1}{\sigma_n}.$$

Therefore, the condition number of A is

$$\|A^{-1}\| \|A\| = \frac{\sigma_1}{\sigma_n}.$$

Activity 34.4. Let $A = \begin{bmatrix} 1.0000 & 1.0000 \\ 1.0000 & 1.0005 \end{bmatrix}$. A computer algebra system gives the singular values of A as 2.00025003124999934 and 0.000249968750000509660. What is the condition number of A . What does that tell us about A ? Does this seem reasonable given the result of Activity 34.3?

Activity 34.5.

- (a) What is the smallest the condition number of a matrix can be? Find an entire class of matrices with this smallest condition number.
- (b) What is the condition number of an orthogonal matrix? Why does this make sense? (Hint: If P is an orthogonal matrix, what is $\|Px\|$ for any vector x ? What does this make $\|P\|$?)
- (c) What is the condition number of an invertible symmetric matrix in terms of its eigenvalues?
- (d) Why do we not define the condition number of a non-invertible matrix? If we did, what would the condition number have to be? Why?

Pseudoinverses

Not every matrix is invertible, so we cannot always solve a matrix equation $Ax = b$. However, every matrix has a pseudoinverse A^+ that acts something like an inverse. Even when we can't solve a matrix equation $Ax = b$ because b isn't in $\text{Col } A$, we can use the pseudoinverse of A to "solve" the equation $Ax = b$ with the "solution" A^+b . While not an exact solution, A^+b turns out to be the best approximation to a solution in the least squares sense. In this use the singular value decomposition to find the pseudoinverse of a matrix.

Preview Activity 34.2. Let $A = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}$. The singular value decomposition of A is $U\Sigma V^\top$



where

$$U = \frac{\sqrt{2}}{2} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix},$$

$$\Sigma = \begin{bmatrix} \sqrt{3} & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix},$$

$$V = \frac{1}{6} \begin{bmatrix} \sqrt{6} & -3\sqrt{2} & 2\sqrt{3} \\ 2\sqrt{6} & 0 & -2\sqrt{3} \\ \sqrt{6} & 3\sqrt{2} & 2\sqrt{3} \end{bmatrix}.$$

- (1) Explain why A is not an invertible matrix.
- (2) Explain why the matrices U and V are invertible. How are U^{-1} and V^{-1} related to U^\top and V^\top ?
- (3) Recall that one property of invertible matrices is that the inverse of a product of invertible matrices is the product of the inverses in the reverse order. If A were invertible, then A^{-1} would be $(U\Sigma V^\top)^{-1} = V\Sigma^{-1}U^\top$. Even though U and V are invertible, the matrix Σ is not. But Σ does contain non-zero eigenvalues that have reciprocals, so consider the matrix $\Sigma^+ = \begin{bmatrix} \frac{1}{\sqrt{3}} & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}$. Calculate the products $\Sigma\Sigma^+$ and $\Sigma^+\Sigma$. How are the results similar to that obtained with a matrix inverse?
- (4) The only matrix in the singular value decomposition of A that is not invertible is Σ . But the matrix Σ^+ acts somewhat like an inverse of Σ , so let us define A^+ as $V\Sigma^+U^\top$. Now we explore a few properties of the matrix A^+ .
 - (a) Calculate AA^+ and A^+A for $A = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}$. What do you notice?
 - (b) Calculate A^+AA^+ and AA^+A for $A = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}$. What do you notice?

Only some square matrices have inverses. However, every matrix has a pseudoinverse. A pseudoinverse A^+ of a matrix A provides something like an inverse when a matrix doesn't have an inverse. Pseudoinverses are useful to approximate solutions to linear systems. If A is invertible, then the equation $Ax = b$ has the solution $x = A^{-1}b$, but when A is not invertible and b is not in $\text{Col } A$, then the equation $Ax = b$ has no solution. In the invertible case of an $n \times n$ matrix A , there is a matrix B so that $AB = BA = I_n$. This also implies that $BAB = B$ and $ABA = A$. To mimic this situation when A is not invertible, we search for a matrix A^+ (a pseudoinverse of A) so that $AA^+A = A$ and $A^+AA^+ = A^+$, as we saw in Preview Activity 34.2. Then it turns out that A^+ acts something like an inverse for A . In this case, we approximate the solution to $Ax = b$ by $x^* = A^+b$, and we will see that the vector $Ax^* = AA^+b$ turns out to be the vector in $\text{Col } A$ that is closest to b in the least squares sense.

A reasonable question to ask is how we can find a pseudoinverse of a matrix A . A singular value decomposition provides an answer to this question. If A is an invertible $n \times n$ matrix, then 0

is not an eigenvalue of A . As a result, in the singular value decomposition $U\Sigma V^T$ of A , the matrix Σ is an invertible matrix (note that U , Σ , and V are all $n \times n$ matrices in this case). So

$$A^{-1} = (U\Sigma V^T)^{-1} = V\Sigma^{-1}U^T,$$

where

$$\Sigma^{-1} = \begin{bmatrix} \frac{1}{\sigma_1} & & & 0 \\ & \frac{1}{\sigma_2} & & \\ & & \frac{1}{\sigma_3} & \\ 0 & & & \ddots \\ & & & \frac{1}{\sigma_n} \end{bmatrix}.$$

In this case, $V\Sigma^{-1}U^T$ is a singular value decomposition for A^{-1} .

To understand in general how a pseudoinverse is found, let A be an $m \times n$ matrix with $m \neq n$, or an $n \times n$ with rank less than n . In these cases A does not have an inverse. But as in Preview Activity 34.2, a singular value decomposition provides a pseudoinverse A^+ for A . Let $U\Sigma V^T$ be a singular value decomposition of an $m \times n$ matrix A of rank r , with

$$\Sigma = \left[\begin{array}{ccc|c} \sigma_1 & & & 0 \\ & \sigma_2 & 0 & \\ & & \sigma_3 & \\ 0 & & & \ddots \\ & & & \sigma_r \\ \hline 0 & & & 0 \end{array} \right]$$

The matrices U and V are invertible, but the matrix Σ is not if A is not invertible. If we let Σ^+ be the $n \times m$ matrix defined by

$$\Sigma^+ = \left[\begin{array}{ccc|c} \frac{1}{\sigma_1} & & & 0 \\ & \frac{1}{\sigma_2} & 0 & \\ & & \frac{1}{\sigma_3} & \\ 0 & & & \ddots \\ & & & \frac{1}{\sigma_r} \\ \hline 0 & & & 0 \end{array} \right],$$

then Σ^+ will act much like an inverse of Σ might. In fact, it is not difficult to see that

$$\Sigma\Sigma^+ = \left[\begin{array}{c|c} I_r & 0 \\ \hline 0 & 0 \end{array} \right] \text{ and } \Sigma^+\Sigma = \left[\begin{array}{c|c} I_r & 0 \\ \hline 0 & 0 \end{array} \right],$$

where $\Sigma\Sigma^+$ is an $m \times m$ matrix and $\Sigma^+\Sigma$ is an $n \times n$ matrix.

The matrix

$$A^+ = V\Sigma^+U^T \tag{34.8}$$

is a *pseudoinverse* of A .

Activity 34.6.



- (a) Find the pseudoinverse A^+ of $A = \begin{bmatrix} 0 & 5 \\ 4 & 3 \\ -2 & 1 \end{bmatrix}$. Use the singular value decomposition $U\Sigma V^\top$ of A , where

$$U = \begin{bmatrix} \frac{\sqrt{2}}{2} & \frac{\sqrt{3}}{3} & 1 \\ \frac{\sqrt{2}}{2} & -\frac{\sqrt{3}}{3} & 0 \\ 0 & \frac{\sqrt{3}}{3} & 0 \end{bmatrix}, \Sigma = \begin{bmatrix} \sqrt{40} & 0 \\ 0 & \sqrt{15} \\ 0 & 0 \end{bmatrix}, V = \frac{1}{\sqrt{5}} \begin{bmatrix} 1 & -2 \\ 2 & 1 \end{bmatrix}.$$

- (b) The vector $\mathbf{b} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$ is not in $\text{Col } A$. The vector $\mathbf{x}^* = A^+ \mathbf{b}$ is an approximation to a solution of $A\mathbf{x} = \mathbf{b}$, and $AA^+ \mathbf{b}$ is in $\text{Col } A$. Find $A\mathbf{x}^*$ and determine how far $A\mathbf{x}^*$ is from \mathbf{b} .

Pseudoinverses satisfy several properties that are similar to those of inverses. For example, we had an example in Preview Activity 34.2 where $AA^+A = A$ and $A^+AA^+ = A^+$. That A^+ always satisfies these properties is the subject of the next activity.

Activity 34.7. Let A be an $m \times n$ matrix with singular value decomposition $U\Sigma V^\top$. Let A^+ be defined as in (34.8).

- (a) Show that $AA^+A = A$.
- (b) Show that $A^+AA^+ = A^+$.

Activity 34.7 shows that A^+ satisfies properties that are similar to those of an inverse of A . In fact, A^+ satisfies several other properties (that together can be used as defining properties) as stated in the next theorem. The conditions of Theorem 34.3 are called the *Penrose* or *Moore-Penrose* conditions.³ Verification of the remaining parts of this theorem are left for the exercises.

Theorem 34.3 (The Moore-Penrose Conditions.). *A pseudoinverse of a matrix A is a matrix A^+ that satisfies the following properties.*

- (1) $AA^+A = A$
- (2) $A^+AA^+ = A^+$
- (3) $(AA^+)^\top = AA^+$
- (4) $(A^+A)^\top = A^+A$

Also, there is a unique matrix A^+ that satisfies these properties. The verification of this property is left to the exercises.

³Theorem 34.3 is often given as the definition of a pseudoinverse.



Least Squares Approximations

In many situations we want to be able to make predictions from data. However, data is rarely well-behaved and so we need to use approximation techniques to estimate from the data.

Preview Activity 34.3. NBC was awarded the U.S. television broadcast rights to the 2016 and 2020 summer Olympic games. Table 34.2 lists the amounts paid (in millions of dollars) by NBC sports for the 2008 through 2012 summer Olympics plus the recently concluded bidding for the 2016 and 2020 Olympics, where year 0 is the year 2008. Figure 34.3 shows a plot of the data. Our goal in this activity is to find a linear function f defined by $f(x) = a_0 + a_1x$ that fits the data well.

Year	Amount
0	894
4	1180
8	1226
12	1418

Table 34.2: Olympics data.

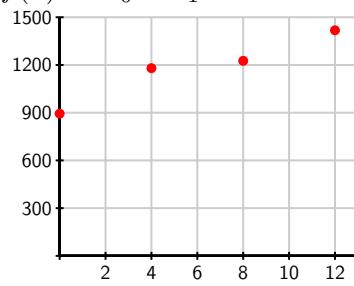


Figure 34.3: A plot of the data.

If the data were actually linear, then the data would satisfy the system

$$a_0 + 0a_1 = 894$$

$$a_0 + 4a_1 = 1180$$

$$a_0 + 8a_1 = 1226$$

$$a_0 + 12a_1 = 1418.$$

In matrix form this system is $Ax = b$, where $A = \begin{bmatrix} 1 & 0 \\ 1 & 4 \\ 1 & 8 \\ 1 & 12 \end{bmatrix}$, $x = \begin{bmatrix} a_0 \\ a_1 \end{bmatrix}$, and $b = \begin{bmatrix} 894 \\ 1180 \\ 1226 \\ 1418 \end{bmatrix}$.

Assume that a singular value decomposition of A (with entries rounded to 4 decimal places) is $U\Sigma V^T$, where

$$U \approx \begin{bmatrix} 0.0071 & -0.8366 & 0.2236 & 0.5000 \\ 0.2713 & -0.4758 & -0.6708 & -0.5000 \\ 0.5355 & -0.1150 & 0.6708 & -0.5000 \\ 0.7997 & 0.2459 & -0.2236 & 0.5000 \end{bmatrix},$$

$$\Sigma \approx \begin{bmatrix} 15.0528 & 0 \\ 0 & 1.1884 \\ 0 & 0 \\ 0 & 0 \end{bmatrix},$$

$$V \approx \begin{bmatrix} 0.1072 & -0.9942 \\ 0.9942 & 0.1072 \end{bmatrix}.$$

- (1) Explain why A is not an invertible matrix. Find a pseudoinverse A^+ of A and calculate A^+b . Round calculations to four decimal places.

- (2) Use A^+b to find the linear approximation $f(x) = a_0 + a_1x$. Plot your linear approximation against the data in Figure 34.3.

Preview Activity 34.3 illustrates how A^+b can be used as an approximation to a solution to the equation $Ax = b$ when b is not in $\text{Col } A$. Now we will examine what kind of approximation A^+b actually is.

Let $U\Sigma V^\top$ be a singular value decomposition for an $m \times n$ matrix A of rank r . Then the columns of

$$U = [\mathbf{u}_1 \ \mathbf{u}_2 \ \cdots \ \mathbf{u}_m]$$

form an orthonormal basis for \mathbb{R}^m and $\{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_r\}$ is a basis for $\text{Col } A$. Remember from Section 30 that if b is any vector in \mathbb{R}^m , then

$$\text{proj}_{\text{Col } A} b = (b \cdot \mathbf{u}_1)\mathbf{u}_1 + (b \cdot \mathbf{u}_2)\mathbf{u}_2 + \cdots + (b \cdot \mathbf{u}_r)\mathbf{u}_r$$

is the least squares approximation of the vector b by a vector in $\text{Col } A$. We can extend this sum to all of columns of U as

$$\text{proj}_{\text{Col } A} b = (b \cdot \mathbf{u}_1)\mathbf{u}_1 + (b \cdot \mathbf{u}_2)\mathbf{u}_2 + \cdots + (b \cdot \mathbf{u}_r)\mathbf{u}_r + 0\mathbf{u}_{r+1} + 0\mathbf{u}_{r+2} + \cdots + 0\mathbf{u}_m.$$

It follows that

$$\begin{aligned} \text{proj}_{\text{Col } A} b &= \sum_{i=1}^r \mathbf{u}_i (\mathbf{u}_i \cdot b) \\ &= \sum_{i=1}^r \mathbf{u}_i (\mathbf{u}_i^\top b) \\ &= \sum_{i=1}^r (\mathbf{u}_i \mathbf{u}_i^\top) b \\ &= \left(\sum_{i=1}^r (1)(\mathbf{u}_i \mathbf{u}_i^\top) \right) b + \left(\sum_{i=r+1}^m 0(\mathbf{u}_i \mathbf{u}_i^\top) \right) b \\ &= (UDU^\top)b, \end{aligned}$$

where

$$D = \left[\begin{array}{c|c} I_r & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{0} \end{array} \right].$$

Now, if $\mathbf{z} = A^+b$, then

$$A\mathbf{z} = (U\Sigma V^\top)(V\Sigma^+ U^\top b) = (U\Sigma\Sigma^+ U^\top) b = (UDU^\top) b = \text{proj}_{\text{Col } A} b,$$

and hence the vector $A\mathbf{z} = AA^+b$ is the vector Ax in $\text{Col } A$ that minimizes $\|Ax - b\|$. Thus, $A\mathbf{z}$ is in actuality the least squares approximation to b . So a singular value decomposition allows us to construct the pseudoinverse of a matrix A and then directly solve the least squares problem.

Activity 34.8. Having to calculate eigenvalues and eigenvectors for a matrix to produce a singular value decomposition to find pseudoinverse can be computationally intense. As we demonstrate in this activity, the process is easier if the columns of A are linearly independent. More specifically, we will prove the following theorem.



Theorem 34.4. If the columns of a matrix A are linearly independent, then $A^+ = (A^\top A)^{-1} A^\top$.

To see how, suppose that A is an $m \times n$ matrix with linearly independent columns.

- (a) Given that the columns of A are linearly independent, what must be the relationship between n and m ?
- (b) Since the columns of A are linearly independent, it follows that $A^\top A$ is invertible (see Exercise 13.). So the eigenvalues of A are all non-zero. Let $\sigma_1, \sigma_2, \dots, \sigma_r$ be the singular values of A . How is r related to n , and what do Σ and Σ^+ look like?
- (c) Let us now investigate the form of the invertible matrix $A^\top A$ (note that neither A nor A^\top is necessarily invertible). If a singular value decomposition of A is $U\Sigma V^\top$, show that

$$A^\top A = V\Sigma^\top \Sigma V^\top.$$

- (d) Let $\lambda_i = \sigma_i^2$ for i from 1 to n . It is straightforward to see that $\Sigma^\top \Sigma$ is an $n \times n$ diagonal matrix D , where

$$D = \Sigma^\top \Sigma = \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & 0 \\ & & \lambda_3 & \\ & & & \ddots \\ 0 & & & \lambda_n \end{bmatrix}.$$

Then $(A^\top A)^{-1} = VD^{-1}V^\top$. Recall that $A^+ = V\Sigma^+U^\top$, so to relate $A^\top A$ to A^+ we need a product that is equal to Σ^+ . Explain why

$$D^{-1}\Sigma^\top = \Sigma^+.$$

- (e) Complete the activity by showing that

$$(A^\top A)^{-1} A^\top = A^+.$$

Therefore, to calculate A^+ and solve a least squares problem, Theorem 34.4 shows that as long as the columns of A are linearly independent, we can avoid using a singular value decomposition

of A in finding A^+ . As an example, if $A = \begin{bmatrix} 1 & 0 \\ 1 & 4 \\ 1 & 8 \\ 1 & 12 \end{bmatrix}$ and $\mathbf{b} = \begin{bmatrix} 894 \\ 1180 \\ 1226 \\ 1418 \end{bmatrix}$ as in Preview Activity 34.3, then

$$(A^\top A)^{-1} A^\top \approx \begin{bmatrix} 0.7000 & 0.4000 & 0.1000 & -0.2000 \\ -0.0750 & -0.0250 & 0.0250 & 0.0750 \end{bmatrix}$$

and

$$(A^\top A)^{-1} A^\top \mathbf{b} \approx \begin{bmatrix} 936.8000 \\ 40.4500 \end{bmatrix}.$$

So $f(x) = 936.8 + 40.45x$ is (to 4 decimal places) the least squares linear approximation to the data. A graph of f versus the data is shown in Figure 34.4. By least squares we mean that $f(x)$ approximates the data so that the sum of the squares of the vertical distances between the data and corresponding values of f (as illustrated in Figure 34.4) is as small as possible.



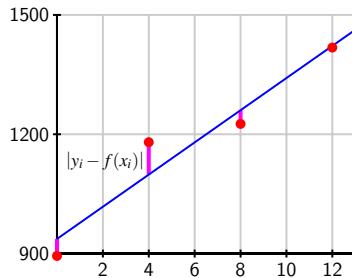


Figure 34.4: Linear approximation plotted against the data set.

Activity 34.9. Suppose we wanted to fit a quadratic or a cubic polynomial to a set of data. Preview Activity 34.3 showed us how to fit a line to data and we can extend that idea to any degree polynomial. To fit a polynomial $p(x) = a_n x^n + a_{n-1} x^{n-1} + \dots + a_1 + a_0$ of degree n to m data points $(x_1, y_1), (x_2, y_2), \dots, (x_m, y_m)$, no two of which have the same x coordinate in the least squares sense, we want to find the least squares approximate solution to the system .

$$\begin{aligned} y_1 &= a_0 + x_1 a_1 + x_1^2 a_2 + \dots + x_1^{n-1} a_{n-1} + x_1^n a_n \\ y_2 &= a_0 + x_2 a_1 + x_2^2 a_2 + \dots + x_2^{n-1} a_{n-1} + x_2^n a_n \\ y_3 &= a_0 + x_3 a_1 + x_3^2 a_2 + \dots + x_3^{n-1} a_{n-1} + x_3^n a_n \\ &\vdots \\ y_m &= a_0 + x_m a_1 + x_m^2 a_2 + \dots + x_m^{n-1} a_{n-1} + x_m^n a_n. \end{aligned}$$

of m equations in the $n + 1$ unknowns a_0, a_1, \dots, a_{n-1} , and a_n . In matrix form we can write this system as $M\mathbf{a} = \mathbf{y}$, where

$$M = \begin{bmatrix} 1 & x_1 & x_1^2 & \cdots & x_1^{n-1} & x_1^n \\ 1 & x_2 & x_2^2 & \cdots & x_2^{n-1} & x_2^n \\ 1 & x_3 & x_3^2 & \cdots & x_3^{n-1} & x_3^n \\ \vdots & \vdots & \vdots & \cdots & \vdots & \vdots \\ 1 & x_m & x_m^2 & \cdots & x_m^{n-1} & x_m^n \end{bmatrix}$$

while the vectors \mathbf{a} and \mathbf{y} are

$$\mathbf{a} = \begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ \vdots \\ a_{n-1} \\ a_n \end{bmatrix} \text{ and } \mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ \vdots \\ y_{m-1} \\ y_m \end{bmatrix}.$$

- (a) Set up the matrix equation to fit a quadratic to the Olympics data in Table Table 34.2.
- (b) Use Theorem 34.4 and appropriate technology to find the least squares quadratic polynomial for the Olympics data. Draw your approximation against the data as shown in Figure 34.3. Round approximations to 4 decimal places.

Examples

What follows are worked examples that use the concepts from this section.

Example 34.5. Let

$$A = \begin{bmatrix} 2 & 5 & 4 \\ 6 & 3 & 0 \\ 6 & 3 & 0 \\ 2 & 5 & 4 \end{bmatrix}.$$

The eigenvalues of $A^T A$ are $\lambda_1 = 144$, $\lambda_2 = 36$, and $\lambda_3 = 0$ with corresponding eigenvectors

$$\mathbf{w}_1 = \begin{bmatrix} 2 \\ 2 \\ 1 \end{bmatrix}, \quad \mathbf{w}_2 = \begin{bmatrix} -2 \\ 1 \\ 2 \end{bmatrix}, \quad \text{and} \quad \mathbf{w}_3 = \begin{bmatrix} 1 \\ -2 \\ 2 \end{bmatrix}.$$

In addition,

$$A\mathbf{w}_1 = \begin{bmatrix} 18 \\ 18 \\ 18 \\ 18 \end{bmatrix} \quad \text{and} \quad A\mathbf{w}_2 = \begin{bmatrix} 9 \\ -9 \\ -9 \\ 9 \end{bmatrix}.$$

- (a) Find orthogonal matrices U and V , and the matrix Σ , so that $U\Sigma V^T$ is a singular value decomposition of A .
- (b) Determine the best rank 1 approximation to A . Give an appropriate numerical estimate as to how good this approximation is to A .
- (c) Find the pseudoinverse A^+ of A .

- (d) Let $\mathbf{b} = \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix}^T$. Does the matrix equation

$$A\mathbf{x} = \mathbf{b}$$

have a solution? If so, find the solution. If not, find the best approximation you can to a solution to this matrix equation.

- (e) Use the orthogonal basis $\{\frac{1}{2}[1 \ 1 \ 1]^T, \frac{1}{2}[1 \ -1 \ -1]^T\}$ of $\text{Col } A$ to find the projection of \mathbf{b} onto $\text{Col } A$. Compare to your solution in part (c).

Example Solution.

- (a) Normalizing the eigenvectors \mathbf{w}_1 , \mathbf{w}_2 , and \mathbf{w}_3 to normal eigenvectors \mathbf{v}_1 , \mathbf{v}_2 , and \mathbf{v}_3 , respectively, gives us an orthogonal matrix

$$V = \begin{bmatrix} \frac{2}{3} & -\frac{2}{3} & \frac{1}{3} \\ \frac{2}{3} & \frac{1}{3} & -\frac{2}{3} \\ \frac{1}{3} & \frac{2}{3} & \frac{2}{3} \end{bmatrix}.$$

Now $A\mathbf{v}_i = A \frac{\mathbf{w}_i}{\|\mathbf{w}_i\|} = \frac{1}{\|\mathbf{w}_i\|} A\mathbf{w}_i$, so normalizing the vectors $A\mathbf{w}_1$ and $A\mathbf{w}_2$ gives us vectors

$$\mathbf{u}_1 = \frac{1}{2} \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \quad \text{and} \quad \mathbf{u}_2 = \frac{1}{2} \begin{bmatrix} 1 \\ -1 \\ -1 \\ 1 \end{bmatrix}$$

that are the first two columns of our matrix U . Given that U is a 4×4 matrix, we need to find two other vectors orthogonal to \mathbf{u}_1 and \mathbf{u}_2 that will combine with \mathbf{u}_1 and \mathbf{u}_2 to form an orthogonal basis for \mathbb{R}^4 . Letting $\mathbf{z}_1 = [1 \ 1 \ 1 \ 1]^\top$, $\mathbf{z}_2 = [1 \ -1 \ -1 \ 1]^\top$, $\mathbf{z}_3 = [1 \ 0 \ 0 \ 0]^\top$, and $\mathbf{z}_4 = [0 \ 1 \ 0 \ 1]^\top$, a computer algebra system shows that the reduced row echelon form of the matrix $[\mathbf{z}_1 \ \mathbf{z}_2 \ \mathbf{z}_3 \ \mathbf{z}_4]$ is I_4 , so that vectors $\mathbf{z}_1, \mathbf{z}_2, \mathbf{z}_3, \mathbf{z}_4$ are linearly independent. Letting $\mathbf{w}_1 = \mathbf{z}_1$ and $\mathbf{w}_2 = \mathbf{z}_2$, the Gram-Schmidt process shows that the set $\{\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3, \mathbf{w}_4\}$ is an orthogonal basis for \mathbb{R}^4 , where $\mathbf{w}_3 = \frac{1}{4}[2 \ 0 \ 0 \ -2]^\top$ and (using $[1 \ 0 \ 0 \ -1]^\top$ for \mathbf{w}_3) $\mathbf{w}_4 = \frac{1}{4}[0 \ 2 \ -2 \ 0]^\top$.

The set $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3, \mathbf{u}_4\}$ where $\mathbf{u}_1 = \frac{1}{2}[1 \ 1 \ 1 \ 1]^\top$, $\mathbf{u}_2 = \frac{1}{2}[1 \ -1 \ -1 \ 1]^\top$, $\mathbf{u}_3 = \frac{1}{\sqrt{2}}[1 \ 0 \ 0 \ -1]^\top$ and $\mathbf{u}_4 = \frac{1}{\sqrt{2}}[0 \ 1 \ -1 \ 0]^\top$ is an orthonormal basis for \mathbb{R}^4 and we can let

$$U = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & \frac{1}{\sqrt{2}} & 0 \\ \frac{1}{2} & -\frac{1}{2} & 0 & \frac{1}{\sqrt{2}} \\ \frac{1}{2} & -\frac{1}{2} & 0 & -\frac{1}{\sqrt{2}} \\ \frac{1}{2} & \frac{1}{2} & -\frac{1}{\sqrt{2}} & 0 \end{bmatrix}.$$

The singular values of A are $\sigma_1 = \sqrt{\lambda_1} = 12$ and $\sigma_2 = \sqrt{\lambda_2} = 6$, and so

$$\Sigma = \begin{bmatrix} 12 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

Therefore, a singular value decomposition of A is $U\Sigma V^\top$ of

$$\begin{bmatrix} \frac{1}{2} & \frac{1}{2} & \frac{1}{\sqrt{2}} & 0 \\ \frac{1}{2} & -\frac{1}{2} & 0 & \frac{1}{\sqrt{2}} \\ \frac{1}{2} & -\frac{1}{2} & 0 & -\frac{1}{\sqrt{2}} \\ \frac{1}{2} & \frac{1}{2} & -\frac{1}{\sqrt{2}} & 0 \end{bmatrix} \begin{bmatrix} 12 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \frac{2}{3} & \frac{2}{3} & \frac{1}{3} \\ -\frac{2}{3} & \frac{1}{3} & \frac{2}{3} \\ \frac{1}{3} & -\frac{2}{3} & \frac{2}{3} \end{bmatrix}.$$

(b) The outer product decomposition of A is

$$A = \sigma_1 \mathbf{u}_1 \mathbf{v}_1^\top + \sigma_2 \mathbf{u}_2 \mathbf{v}_2^\top.$$

So the rank one approximation to A is

$$\sigma_1 \mathbf{u}_1 \mathbf{v}_1^\top = 12 \left(\frac{1}{2} \right) \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \begin{bmatrix} \frac{2}{3} & \frac{2}{3} & \frac{1}{3} \end{bmatrix} = \begin{bmatrix} 4 & 4 & 2 \\ 4 & 4 & 2 \\ 4 & 4 & 2 \\ 4 & 4 & 2 \end{bmatrix}.$$



The error in approximating A with this rank one approximation is

$$\sqrt{\frac{\sigma_2^2}{\sigma_1^2 + \sigma_2^2}} = \sqrt{\frac{36}{180}} = \sqrt{\frac{1}{5}} \approx 0.447.$$

- (c) Given that $A = U\Sigma V^\top$, we use the pseudoinverse Σ^+ of Σ to find the pseudoinverse A^+ of A by

$$A^+ = V\Sigma^+U^\top.$$

Now

$$\Sigma^+ = \begin{bmatrix} \frac{1}{12} & 0 & 0 \\ 0 & \frac{1}{6} & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

so

$$\begin{aligned} A^+ &= \begin{bmatrix} \frac{2}{3} & -\frac{2}{3} & \frac{1}{3} \\ \frac{2}{3} & \frac{1}{3} & -\frac{2}{3} \\ \frac{1}{3} & \frac{2}{3} & \frac{2}{3} \end{bmatrix} \begin{bmatrix} \frac{1}{12} & 0 & 0 \\ 0 & \frac{1}{6} & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & \frac{1}{\sqrt{2}} & 0 \\ \frac{1}{2} & -\frac{1}{2} & 0 & \frac{1}{\sqrt{2}} \\ \frac{1}{2} & -\frac{1}{2} & 0 & -\frac{1}{\sqrt{2}} \\ \frac{1}{2} & \frac{1}{2} & -\frac{1}{\sqrt{2}} & 0 \end{bmatrix}^\top \\ &= \frac{1}{72} \begin{bmatrix} -2 & 6 & 6 & -2 \\ 4 & 0 & 0 & 4 \\ 5 & -3 & -3 & 5 \end{bmatrix}. \end{aligned}$$

- (d) Augmenting A with \mathbf{b} and row reducing shows that

$$[A \ \mathbf{b}] \sim \begin{bmatrix} 2 & 5 & 4 & 1 \\ 0 & -12 & -12 & -3 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix},$$

so \mathbf{b} is not in $\text{Col } A$ and the equation $A\mathbf{x} = \mathbf{b}$ has no solution. However, the best approximation to a solution to $A\mathbf{x} = \mathbf{b}$ is found using the pseudoinverse A^+ of A . That best solution is

$$\begin{aligned} \mathbf{x}^* &= AA^+\mathbf{b} \\ &= \begin{bmatrix} 2 & 5 & 4 \\ 6 & 3 & 0 \\ 6 & 3 & 0 \\ 2 & 5 & 4 \end{bmatrix} \frac{1}{72} \begin{bmatrix} -2 & 6 & 6 & -2 \\ 4 & 0 & 0 & 4 \\ 5 & -3 & -3 & 5 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix} \\ &= \frac{1}{2} \begin{bmatrix} 2 \\ 1 \\ 1 \\ 2 \end{bmatrix}. \end{aligned}$$

- (e) The rank of A is 2 and an orthonormal basis for $\text{Col } A$ is $\{\mathbf{u}_1, \mathbf{u}_2\}$, where $\mathbf{u}_1 = \frac{1}{2}[1 \ 1 \ 1 \ 1]^\top$ and $\mathbf{u}_2 = \frac{1}{2}[1 \ -1 \ -1 \ 1]^\top$. So

$$\begin{aligned}\text{proj}_{\text{Col } A} \mathbf{b} &= (\mathbf{b} \cdot \mathbf{u}_1)\mathbf{u}_1 + (\mathbf{b} \cdot \mathbf{u}_2)\mathbf{u}_2 \\ &= \left(\frac{3}{2}\right)\left(\frac{1}{2}\right)[1 \ 1 \ 1 \ 1]^\top + \left(\frac{1}{2}\right)\left(\frac{1}{2}\right)[1 \ -1 \ -1 \ 1]^\top \\ &= \frac{1}{2}[2 \ 1 \ 1 \ 2]^\top\end{aligned}$$

as expected from part (c).

Example 34.6. According to the Centers for Disease Control and Prevention⁴, the average length of a male infant (in centimeters) in the US as it ages (with time in months from 1.5 to 8.5) is given in Table 34.3. In this problem we will find the line and the quadratic of best fit in the least squares

Age (months)	1.5	2.5	3.5	4.5	5.5	6.5	7.5	8.5
Average Length (cm)	56.6	59.6	62.1	64.2	66.1	67.9	69.5	70.9

Table 34.3: Average lengths of male infants.

sense to this data. We treat time in months as the independent variable and length in centimeters as the dependent variable.

- (a) Find a line that is the best fit to the data in the least squares sense. Draw a picture of your least squares solution against a scatterplot of the data.
- (b) Now find the least squares quadratic of the form $q(x) = a_2x^2 + a_1x + a_0$ to the data. Draw a picture of your least squares solution against a scatterplot of the data.

Example Solution.

- (a) We assume that a line of the form $f(x) = a_1x + a_0$ contains all of the data points. The first date point would satisfy $1.5a_1 + a_0 = 56.6$, the second $2.5a_1 + a_0 = 59.6$, and so on, giving us the linear system

$$\begin{aligned}1.5a_1 + a_0 &= 56.6 \\ 2.5a_1 + a_0 &= 59.6 \\ 3.5a_1 + a_0 &= 62.1 \\ 4.5a_1 + a_0 &= 64.2 \\ 5.5a_1 + a_0 &= 66.1 \\ 6.5a_1 + a_0 &= 67.9 \\ 7.5a_1 + a_0 &= 69.5 \\ 8.5a_1 + a_0 &= 70.9.\end{aligned}$$

⁴ https://www.cdc.gov/growthcharts/html_charts/lenageinf.htm



Letting

$$A = \begin{bmatrix} 1.5 & 1 \\ 2.5 & 1 \\ 3.5 & 1 \\ 4.5 & 1 \\ 5.5 & 1 \\ 6.5 & 1 \\ 7.5 & 1 \\ 8.5 & 1 \end{bmatrix}, \quad \mathbf{x} = \begin{bmatrix} a_1 \\ a_0 \end{bmatrix}, \quad \text{and} \quad \mathbf{b} = \begin{bmatrix} 56.6 \\ 59.6 \\ 62.1 \\ 64.2 \\ 66.1 \\ 67.9 \\ 69.5 \\ 70.9 \end{bmatrix},$$

we can write this system in the matrix form $A\mathbf{x} = \mathbf{b}$. Neither column of A is a multiple of the other, so the columns of A are linearly independent. The least squares solution to the system is then found by

$$A^+ \mathbf{b} = (A^\top A)^{-1} A^\top \mathbf{b}.$$

Technology shows that (with entries rounded to 3 decimal places), A^+ is

$$\begin{bmatrix} -0.083 & -0.060 & -0.036 & -0.012 & 0.012 & 0.036 & 0.060 & 0.083 \\ 0.542 & 0.423 & 0.304 & 0.185 & 0.065 & -0.054 & -0.173 & -0.292 \end{bmatrix},$$

and

$$A^+ \mathbf{b} \approx \begin{bmatrix} 2.011 \\ 54.559 \end{bmatrix}.$$

So the least squares linear solution to $A\mathbf{x} = \mathbf{b}$ is f defined by $f(x) \approx 2.011x + 54.559$. A graph of f against the data points is shown at left in Figure 34.5.

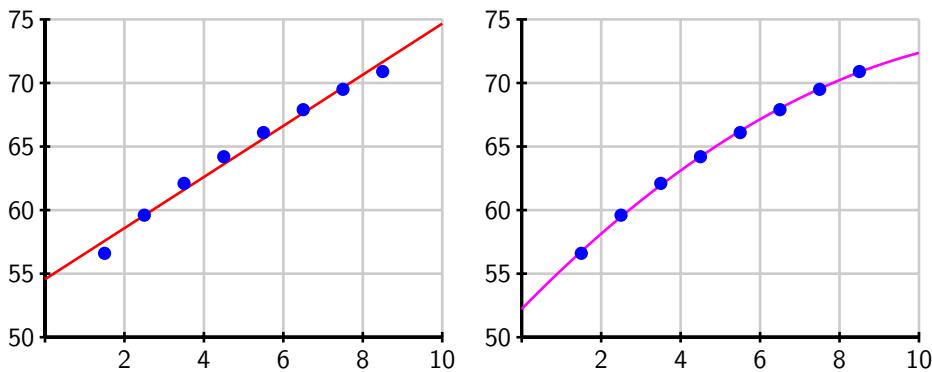


Figure 34.5: Left: Least squares line. Right: Least squares quadratic.

- (b) The first date point would satisfy $(1.5^2)a_2 + 1.5a_1 + a_0 = 56.6$, the second $(2.5^2)a_2 +$

$2.5a_1 + a_0 = 59.6$, and so on, giving us the linear system

$$\begin{aligned} 1.5^2 a_2 + 1.5 a_1 + a_0 &= 56.6 \\ 2.5^2 a_2 + 2.5 a_1 + a_0 &= 59.6 \\ 3.5^2 a_2 + 3.5 a_1 + a_0 &= 62.1 \\ 4.5^2 a_2 + 4.5 a_1 + a_0 &= 64.2 \\ 5.5^2 a_2 + 5.5 a_1 + a_0 &= 66.1 \\ 6.5^2 a_2 + 6.5 a_1 + a_0 &= 67.9 \\ 7.5^2 a_2 + 7.5 a_1 + a_0 &= 69.5 \\ 8.5^2 a_2 + 8.5 a_1 + a_0 &= 70.9. \end{aligned}$$

Letting

$$A = \begin{bmatrix} 1.5^2 & 1.5 & 1 \\ 2.5^2 & 2.5 & 1 \\ 3.5^2 & 3.5 & 1 \\ 4.5^2 & 4.5 & 1 \\ 5.5^2 & 5.5 & 1 \\ 6.5^2 & 6.5 & 1 \\ 7.5^2 & 7.5 & 1 \\ 8.5^2 & 8.5 & 1 \end{bmatrix}, \quad \mathbf{x} = \begin{bmatrix} a_2 \\ a_1 \\ a_0 \end{bmatrix}, \quad \text{and} \quad \mathbf{b} = \begin{bmatrix} 56.6 \\ 59.6 \\ 62.1 \\ 64.2 \\ 66.1 \\ 67.9 \\ 69.5 \\ 70.9 \end{bmatrix},$$

we can write this system in the matrix form $A\mathbf{x} = \mathbf{b}$.

Technology shows that every column of the reduced row echelon form of A contains a pivot, so the columns of A are linearly independent. The least squares solution to the system is then found by

$$A^+ \mathbf{b} = (A^\top A)^{-1} A^\top \mathbf{b}.$$

Technology shows that (with entries rounded to 3 decimal places) A^+ is

$$\begin{bmatrix} 0.042 & 0.006 & -0.018 & -0.030 & -0.030 & -0.018 & 0.006 & 0.042 \\ -0.500 & -0.119 & 0.143 & 0.286 & 0.310 & 0.214 & 0.000 & -0.333 \\ 1.365 & 0.540 & -0.049 & -0.403 & -0.522 & -0.406 & -0.055 & 0.531 \end{bmatrix},$$

and

$$A^+ \mathbf{b} \approx \begin{bmatrix} -0.118 \\ 3.195 \\ 52.219 \end{bmatrix}.$$

So the least squares quadratic solution to $A\mathbf{x} = \mathbf{b}$ is q defined by $q(x) \approx -0.118x^2 + 3.195x + 52.219$. A graph of q against the data points is shown at right in Figure 34.5.

Summary

- The condition number of an $m \times n$ matrix A is the number $\|A^{-1}\| \|A\|$. The condition number provides a measure of how well the relative error in a calculated value $\Delta \mathbf{b}$ predicts the relative error in $\Delta \mathbf{x}$ when we are trying to solve a system $A\mathbf{x} = \mathbf{b}$.



- A pseudoinverse A^+ of a matrix A can be found through a singular value decomposition. Let $U\Sigma V^\top$ be a singular value decomposition of an $m \times n$ matrix A of rank r , with

$$\Sigma = \begin{bmatrix} \sigma_1 & & & & \\ & \sigma_2 & & 0 & \\ & & \sigma_3 & & \\ & 0 & & \ddots & \\ & & & & \sigma_r \\ \hline 0 & & & & 0 \end{bmatrix}$$

If Σ^+ is the $n \times m$ matrix defined by

$$\Sigma^+ = \begin{bmatrix} \frac{1}{\sigma_1} & & & & \\ & \frac{1}{\sigma_2} & & 0 & \\ & & \frac{1}{\sigma_3} & & \\ & 0 & & \ddots & \\ & & & & \frac{1}{\sigma_r} \\ \hline 0 & & & & 0 \end{bmatrix},$$

then $A^+ = V\Sigma^+U^\top$.

- A pseudoinverse A^+ of a matrix A acts like an inverse for A . So if we can't solve a matrix equation $Ax = b$ because b isn't in $\text{Col } A$, we can use the pseudoinverse of A to "solve" the equation $Ax = b$ with the "solution" A^+b . While not an exact solution, A^+b turns out to be the best approximation to a solution in the least squares sense.
- If the columns of A are linearly independent, then we can alternatively find the least squares solution as $(A^\top A)^{-1}A^\top b$.

Exercises

- (1) Let $A = \begin{bmatrix} 20 & 4 & 32 \\ -4 & 4 & 2 \\ 35 & 22 & 26 \end{bmatrix}$. Then A has singular value decomposition $U\Sigma V^\top$, where

$$U = \frac{1}{5} \begin{bmatrix} 3 & 4 & 0 \\ 0 & 0 & 5 \\ 4 & -3 & 0 \end{bmatrix}$$

$$\Sigma = \begin{bmatrix} 60 & 0 & 0 \\ 0 & 15 & 0 \\ 0 & 0 & 6 \end{bmatrix}$$

$$V = \frac{1}{3} \begin{bmatrix} 2 & -1 & -2 \\ 1 & -2 & 2 \\ 2 & 2 & 1 \end{bmatrix}.$$

- (a) What are the singular values of A ?
- (b) Write the outer product decomposition of A .
- (c) Find the best rank 1 approximation to A . What is the relative error in approximating A by this rank 1 matrix?
- (d) Find the best rank 2 approximation to A . What is the relative error in approximating A by this rank 2 matrix?

$$(2) \text{ Let } A = \begin{bmatrix} 861 & 3969 & 70 & 140 \\ 3969 & 861 & 70 & 140 \\ 3969 & 861 & -70 & -140 \\ 861 & 3969 & -70 & -140 \end{bmatrix}.$$

- (a) Find a singular value decomposition for A .
 - (b) What are the singular values of A ?
 - (c) Write the outer product decomposition of A .
 - (d) Find the best rank 1, 2, and 3 approximations to A . How much information about A does each of these approximations contain?
- (3) The University of Denver Infant Study Center investigated whether babies take longer to learn to crawl in cold months, when they are often bundled in clothes that restrict their movement, than in warmer months. The study sought a relationship between babies' first crawling age and the average temperature during the month they first try to crawl (about 6 months after birth). Some of the data from the study is in Table 34.4. Let x represent the temperature in degrees Fahrenheit and $C(x)$ the average crawling age in months.

x	33	37	48	57
$C(x)$	33.83	33.35	33.38	32.32

Table 34.4: Crawling age.

- (a) Use Theorem 34.4 to find the least squares line to fit this data. Plot the data and your line on the same set of axes.
 - (b) Use your least squares line to predict the average crawling age when the temperature is 65.
- (4) The cost, in cents, of a first class postage stamp in years from 1981 to 1995 is shown in Table 34.5.

Year	1981	1985	1988	1991	1995
Cost	20	22	25	29	32

Table 34.5: Cost of postage.

- (a) Use Theorem 34.4 to find the least squares line to fit this data. Plot the data and your line on the same set of axes.
- (b) Now find the least squares quadratic approximation to this data. Plot the quadratic function on same axes as your linear function.
- (c) Use your least squares line and quadratic to predict the cost of a postage stamp in this year. Look up the cost of a stamp today and determine how accurate your prediction is. Which function gives a better approximation? Provide reasons for any discrepancies.
- (5) Assume that the number of feet traveled by a batted baseball at various angles in degrees (all hit at the same bat speed) is given in Table 34.6.

Angle	10°	20°	30°	40°	50°	60°
Distance	116	190	254	285	270	230

Table 34.6: Distance traveled by batted ball.

- (a) Plot the data and explain why a quadratic function is likely a better fit to the data than a linear function.
- (b) Find the least squares quadratic approximation to this data. Plot the quadratic function on same axes as your data.
- (c) At what angle (or angles), to the nearest degree, must a player bat the ball in order for the ball to travel a distance of 220 feet?
- (6) Not all data is well modeled with polynomials – populations tend to grow at rates proportional to the population, which implies exponential growth. For example, Table 34.7 shows the approximate population of the United States in years between 1920 and 2000, with the population measured in millions. If we assume the population grows exponentially, we would

Year	1920	1930	1940	1950	1960	1970	1980	1990	2000
Population	106	123	142	161	189	213	237	259	291

Table 34.7: U.S. population.

want to find the best fit function f of the form $f(t) = ae^{kt}$, where a and k are constants. To apply the methods we have developed, we could instead apply the natural logarithm to both sides of $y = ae^{kt}$ to obtain the equation $\ln(y) = \ln(a) + kt$. We can then find the best fit line to the data in the form $(t, \ln(y))$ to determine the values of $\ln(a)$ and k . Use this approach to find the best fit exponential function in the least squares sense to the U.S. population data.

- (7) How close can a matrix be to being non-invertible? We explore that idea in this exercise. Let $A = [a_{ij}]$ be the $n \times n$ upper triangular matrix with 1s along the diagonal and with every other entry being -1 .
- (a) What is $\det(A)$? What are the eigenvalues of A ? Is A invertible?



- (b) Let $B = [b_{ij}]$ be the $n \times n$ matrix so that $b_{n1} = -\frac{1}{2^{n-2}}$ and $b_{ij} = a_{ij}$ for all other i and j .
- For the matrix B with $n = 3$, show that the equation $Bx = \mathbf{0}$ has a non-trivial solution. Find one non-trivial solution.
 - For the matrix B with $n = 4$, show that the equation $Bx = \mathbf{0}$ has a non-trivial solution. Find one non-trivial solution.
 - Use the pattern established in parts (i.) and (ii.) to find a non-trivial solution to the equation $Bx = \mathbf{0}$ for an arbitrary value of n . Be sure to verify that you have a solution. Is B invertible? (Hint: For any positive integer m , the sum $1 + \sum_{k=0}^{m-1} 2^k$ is the partial sum of a geometric series with ratio 2 and so $1 + \sum_{k=0}^{m-1} 2^k = 1 + \frac{1-2^m}{1-2} = 2^m$.)
 - Explain why B is not an invertible matrix. Notice that A and B differ by a single entry, and that A is invertible and B is not. Let us examine how close A is to B . Calculate $\|A - B\|_F$? What happens to $\|A - B\|_F$ as n goes to infinity? How close can an invertible matrix be to becoming non-invertible?

(8) Let $A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & -1 & 1 \end{bmatrix}$. In this exercise we find a matrix B so that $B^2 = A$, that is, find a square root of the matrix A .

- Find the eigenvalues and corresponding eigenvectors for A and $A^T A$. Explain what you see.
- Find a matrix V that orthogonally diagonalizes $A^T A$.
- Exercise 8 in Section 33 shows if $U\Sigma V^T$ is a singular value decomposition for a symmetric matrix A , then so is $V\Sigma V^T$. Recall that $A^n = (V\Sigma V^T)^n = V\Sigma^n V^T$ for any positive integer n . We can exploit this idea to define \sqrt{A} to be the matrix

$$V\Sigma^{1/2}V^T,$$

where $\Sigma^{1/2}$ is the matrix whose diagonal entries are the square roots of the corresponding entries of Σ . Let $B = \sqrt{A}$. Calculate B and show that $B^2 = A$.

- Why was it important that A be a symmetric matrix for this process to work, and what had to be true about the eigenvalues of A for this to work?
 - Can you extend the process in this exercise to find a cube root of A ?
- (9) Let A be an $m \times n$ matrix with singular value decomposition $U\Sigma V^T$. Let A^+ be defined as in (34.8). In this exercise we prove the remaining parts of Theorem 34.3.
- Show that $(AA^+)^T = AA^+$. (Hint: $\Sigma\Sigma^+$ is a symmetric matrix.)
 - Show that $(A^+A)^T = A^+A$.
- (10) In this exercise we show that the pseudoinverse of a matrix is the unique matrix that satisfies the Moore-Penrose conditions. Let A be an $m \times n$ matrix with singular value decomposition $U\Sigma V^T$ and pseudoinverse $X = V\Sigma^+U^T$. To show that A^+ is the unique matrix that satisfies the Moore-Penrose conditions, suppose that there is another matrix Y that also satisfies the Moore-Penrose conditions.

- (a) Show that $X = YAX$.
 - (b) Show that $Y = YAX$.
 - (c) How do the results of parts (a) and (b) show that A^+ is the unique matrix satisfying the Moore-Penrose conditions?
- (11) Find the pseudo-inverse of the $m \times n$ zero matrix $A = 0$. Explain the conclusion.
- (12) In all of the examples that we have done finding a singular value decomposition of a matrix, it has been the case (though we haven't mentioned it), that if A is an $m \times n$ matrix, then $\text{rank}(A) = \text{rank}(A^T A)$. Prove this result.
- (13) Show that if the columns of a matrix A are linearly independent, then $A^T A$ is invertible. (Hint: If $A^T A \mathbf{x} = \mathbf{0}$, what is $\mathbf{x}^T A^T A \mathbf{x}$?)
- (14) Label each of the following statements as True or False. Provide justification for your response.
- (a) **True/False** A matrix has a pseudo-inverse if and only if the matrix is singular.
 - (b) **True/False** The pseudoinverse of an invertible matrix A is the matrix A^{-1} .
 - (c) **True/False** If the columns of A are linearly dependent, then there is no least squares solution to $A\mathbf{x} = \mathbf{b}$.
 - (d) **True/False** If the columns of A are linearly independent, then there is a unique least squares solution to $A\mathbf{x} = \mathbf{b}$.
 - (e) **True/False** If T is the matrix transformation defined by a matrix A and S is the matrix transformation defined by A^+ , then T and S are inverse transformations.

Project: GPS and Least Squares

In this project we discuss some of the details about how the GPS works. The idea is based on intersections of spheres. To build a basic understanding of the system, we begin with a 2-dimensional example.

Project Activity 34.1. Suppose that there are three base stations A , B , and C in \mathbb{R}^2 that can send and receive signals from your mobile phone. Assume that A is located at point $(-1, -2)$, B at point $(36, 5)$, and C at point $(16, 35)$. Also assume that your mobile phone location is point (x, y) . Based on the time that it takes to receive the signals from the three base stations, it can be determined that your distance to base station A is 28 km, your distance to base station B is 26 km, and your distance to base station C is 14 km using a coordinate system with measurements in kilometers based on a reference point chosen to be $(0, 0)$. Due to limitations on the measurement equipment, these measurements all contain some unknown error which we will denote as z . The goal is to determine your location in \mathbb{R}^2 based on this information.

If the distance readings were accurate, then the point (x, y) would lie on the circle centered at A of radius 29. The distance from (x, y) to base station A can be represented in two different ways:



28 km and $\sqrt{(x+1)^2 + (y+2)^2}$. However, there is some error in the measurements (due to the receiver clock and satellite clocks not being synchronized), so we really have

$$\sqrt{(x+1)^2 + (y+2)^2} + z = 28,$$

where z is the error. Similarly, (x, y) must also satisfy

$$\sqrt{(x-36)^2 + (y-5)^2} + z = 26$$

and

$$\sqrt{(x-16)^2 + (y-35)^2} + z = 14.$$

- (a) Explain how these three equations can be written in the equivalent form

$$(x+1)^2 + (y+2)^2 = (28-z)^2 \quad (34.9)$$

$$(x-36)^2 + (y-5)^2 = (26-z)^2 \quad (34.10)$$

$$(x-16)^2 + (y-35)^2 = (14-z)^2. \quad (34.11)$$

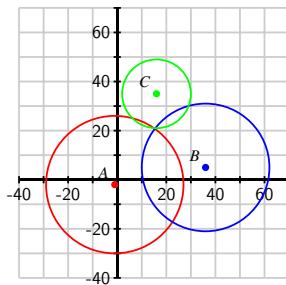


Figure 34.6: Intersections of circles.

- (b) If all measurements were accurate, your position would be at the intersection of the circles centered at A with radius 28 km, centered at B with radius 26 km, and centered at C with radius 14 km as shown in Figure 34.6. Even though the figure might seem to imply it, because of the error in the measurements the three circles do not intersect in one point. So instead, we want to find the best estimate of a point of intersection that we can. The system of equations 34.9, 34.10, and 34.11 is non-linear and can be difficult to solve, if it even has a solution. To approximate a solution, we can linearize the system. To do this, show that if we subtract corresponding sides of equation (34.9) from (34.10) and expand both sides, we can obtain the linear equation

$$37x + 7y + 2z = 712$$

in the unknowns x , y , and z .

- (c) Repeat the process in part (b), subtracting (34.9) from (34.11) and show that we can obtain the linear equation

$$17x + 37y + 14z = 1032$$

in x , y , and z .

(d) We have reduced our system of three non-linear equations to the system

$$\begin{aligned} 37x + 7y + 2z &= 712 \\ 17x + 37y + 14z &= 1032 \end{aligned}$$

of two linear equations in the unknowns x , y , and z . Use technology to find a pseudoinverse of the coefficient matrix of this system. Use the pseudoinverse to find the least squares solution to this system. Does your solution correspond to an approximate point of intersection of the three circles?

Project Activity 34.1 provides the basic idea behind GPS. Suppose you receive a signal from a GPS satellite. The transmission from satellite i provides four pieces of information – a location (x_i, y_i, z_i) and a time stamp t_i according to the satellite's atomic clock. The time stamp allows the calculation of the distance between you and the i th satellite. The transmission travel time is calculated by subtracting the current time on the GPS receiver from the satellite's time stamp. Distance is then found by multiplying the transmission travel time by the rate, which is the speed of light $c = 299792.458$ km/s.⁵ So distance is found as $c(t_i - d)$, where d is the time at the receiver. This signal places your location within in a sphere of that radius from the center of the satellite. If you receive a signal at the same time from two satellites, then your position is at the intersection of two spheres. As can be seen at left in Figure 34.7, that intersection is a circle. So your position has been narrowed quite a bit. Now if you receive simultaneous signals from three spheres, your position is narrowed to the intersection of three spheres, or two points as shown at right in Figure 34.7. So if we could receive perfect information from three satellites, then your location would be exactly determined.

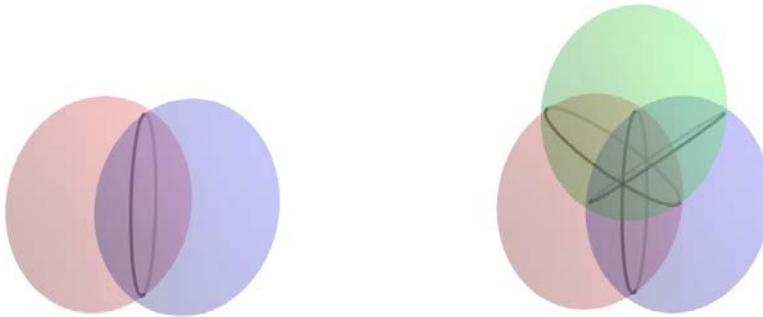


Figure 34.7: Intersections of spheres.

There is a problem with the above analysis – calculating the distances. These distances are determined by the time it takes for the signal to travel from the satellite to the GPS receiver. The times are measured by the clocks in the satellites and the clocks in the receivers. Since the GPS receiver clock is unlikely to be perfectly synchronized with the satellite clock, the distance calculations are not perfect. In addition, the rate at which the signal travels can change as the signal moves through

⁵The signals travel in radio waves, which are electromagnetic waves, and travel at the speed of light. Also, c is the speed of light in a vacuum, but atmosphere is not too dense so we assume this value of c .

the ionosphere and the troposphere. As a result, the calculated distance measurements are not exact, and are referred to as *pseudoranges*. In our calculations we need to factor in the error related to the time discrepancy and other factors. We will incorporate these errors into our measure of d and treat d as an unknown. (Of course, this is all more complicated than is presented here, but this provides the general idea.)

To ensure accuracy, the GPS uses signals from four satellites. Assume a satellite is positioned at point (x_1, y_1, z_1) at a distance d_1 from the GPS receiver located at point (x, y, z) . The distance can also be measured in two ways: as

$$\sqrt{(x - x_1)^2 + (y - y_1)^2 + (z - z_1)^2}.$$

and as $c(t_1 - d)$. So

$$c(t_1 - d) = \sqrt{(x - x_1)^2 + (y - y_1)^2 + (z - z_1)^2}.$$

Again, we are treating d as an unknown, so this equation has the four unknowns x, y, z , and d . Using signals from four satellites produces the system of equations

$$\sqrt{(x - x_1)^2 + (y - y_1)^2 + (z - z_1)^2} = c(t_1 - d) \quad (34.12)$$

$$\sqrt{(x - x_2)^2 + (y - y_2)^2 + (z - z_2)^2} = c(t_2 - d) \quad (34.13)$$

$$\sqrt{(x - x_3)^2 + (y - y_3)^2 + (z - z_3)^2} = c(t_3 - d) \quad (34.14)$$

$$\sqrt{(x - x_4)^2 + (y - y_4)^2 + (z - z_4)^2} = c(t_4 - d). \quad (34.15)$$

Project Activity 34.2. The system of equations (34.12), (34.13), (34.14), and (34.15) is a nonlinear system and is difficult to solve, if it even has a solution. We want a method that will provide at least an approximate solution as well as apply if we use more than four satellites. We choose a reference node (say (x_1, y_1, z_1)) and make calculations relative to that node as we did in Project Activity 34.1.

- (a) First square both sides of the equations (34.12), (34.13), (34.14), and (34.15) to remove the roots. Then subtract corresponding sides of the new first equation (involving (x_1, y_1, z_1)) from the new second equation (involving (x_2, y_2, z_2)) to show that we can obtain the linear equation

$$2(x_2 - x_1)x + 2(y_2 - y_1)y + 2(z_2 - z_1)z + 2c^2(t_1 - t_2)d = c^2(t_1^2 - t_2^2) - h_1 + h_2,$$

where $h_i = x_i^2 + y_i^2 + z_i^2$. (Note that the unknowns are x, y, z , and d – all other quantities are known.)

- (b) Use the result of part (a) to write a linear system that can be obtained by subtracting the first equation from the third and fourth equations as well.
- (c) The linearizations from part (b) determine a system $A\mathbf{x} = \mathbf{b}$ of linear equations. Identify A , \mathbf{x} , and \mathbf{b} . Then explain how we can approximate a best solution to this system in the least squares sense.

We conclude this project with a final note. At times a GPS receiver may only be able to receive signals from three satellites. In these situations, the receiver can substitute the surface of the Earth as a fourth sphere and continue the computation.





Part VIII

Linear Transformations

Section 35

Linear Transformations

Focus Questions

By the end of this section, you should be able to give precise and thorough answers to the questions listed below. You may want to keep these questions in mind to focus your thoughts as you complete the section.

- What is a linear transformation?
- What is the kernel of a linear transformation? What algebraic structure does a kernel of a linear transformation have?
- What is a one-to-one linear transformation? How does its kernel tell us if a linear transformation is one-to-one?
- What is the range of a linear transformation? What algebraic property does the range of a linear transformation possess?
- What is an onto linear transformation? What relationship is there between the codomain and range if a linear transformation is onto?
- What is an isomorphism of vector spaces?

Application: Fractals

Sierpinski triangles and Koch's curves have become common phrases in many mathematics departments across the country. These objects are examples of what are called *fractals*, beautiful geometric constructions that exhibit self-similarity. Fractals are applied in a variety of ways: they help make smaller antennas (e.g., for cell phones) and are used in fiberoptic cables, among other things. In addition, fractals can be used to model realistic objects, such as the Black Spleenwort fern depicted as a fractal image in Figure 35.1. As we will see later in this section, one way to construct a fractal is with an *Iterated Function System* (IFS).



Figure 35.1: An approximation of the Black Spleenwort fern.

Introduction

We have encountered functions throughout our study of mathematics – we explore graphs of functions in algebra and differentiate and integrate functions in calculus. In linear algebra we have investigated special types of functions, e.g., matrix and coordinate transformations, that preserve the vector space structure. Any function that has the same properties as matrix and coordinate transformations is a *linear transformation*. Linear transformations are important in linear algebra in that we can study similarities and connections between vector spaces by examining transformations between them. Linear transformations model or approximately model certain real-life processes (like discrete dynamical systems, geometrical transformations, Google PageRank, etc.). Also, we can determine the behavior of an entire linear transformation by knowing how it acts on just a basis.

Definition 35.1. A **linear transformation** from a vector space V to a vector space W is a function $T : V \rightarrow W$ such that

- (1) $T(\mathbf{u} + \mathbf{v}) = T(\mathbf{u}) + T(\mathbf{v})$ and
- (2) $T(c\mathbf{v}) = cT(\mathbf{v})$

for all \mathbf{u}, \mathbf{v} in V and all scalars c .

These transformations are called *linear* because they respect linear combinations. We can combine both parts of this definition into one statement and say that a mapping T from a vector space V to a vector space W is a linear transformation if

$$T(a\mathbf{u} + b\mathbf{v}) = aT(\mathbf{u}) + bT(\mathbf{v})$$

for all vectors \mathbf{u}, \mathbf{v} in V and all scalars a and b . We can extend this property of a linear transformation (by mathematical induction) to any finite linear combination of vectors. That is, if $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k$ are any vectors in the domain of a linear transformation T and c_1, c_2, \dots, c_k are any scalars, then

$$T(c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_k\mathbf{v}_k) = c_1T(\mathbf{v}_1) + c_2T(\mathbf{v}_2) + \cdots + c_kT(\mathbf{v}_k).$$

This is the property that T respects linear combinations.

Preview Activity 35.1.



- (1) Consider the transformation $T : \mathbb{R}^3 \rightarrow \mathbb{R}^2$ defined by

$$T \left(\begin{bmatrix} x \\ y \\ z \end{bmatrix} \right) = \begin{bmatrix} x + y \\ z + 3 \end{bmatrix}$$

Check that T is not linear by finding two vectors \mathbf{u}, \mathbf{v} which violate the additive property of linear transformations.

- (2) Consider the transformation $T : \mathbb{P}_2 \rightarrow \mathbb{P}_3$ defined by

$$T(a_0 + a_1t + a_2t^2) = (a_0 + a_1) + a_2t + a_1t^3$$

Check that T is a linear transformation.

- (3) Let \mathbb{D} be the set of all differentiable functions from \mathbb{R} to \mathbb{R} . Since $\frac{d}{dx}(0) = 0$, $(f + g)'(x) = f'(x) + g'(x)$ and $(cf)'(x) = cf'(x)$ for any differentiable functions f and g and any scalar c , it follows that \mathbb{D} is a subspace of \mathcal{F} , the vector space of all functions from \mathbb{R} to \mathbb{R} . Let $T : \mathbb{D} \rightarrow \mathcal{F}$ be defined by $T(f) = f'$. Check that T is a linear transformation.
- (4) Every matrix transformation is a linear transformation, so we might expect that general linear transformations share some of the properties of matrix transformations. Let T be a linear transformation from a vector space V to a vector space W . Use the linearity properties to show that $T(\mathbf{0}_V) = \mathbf{0}_W$, where $\mathbf{0}_V$ is the additive identity in V and $\mathbf{0}_W$ is the additive identity in W . (Hint: $\mathbf{0}_V + \mathbf{0}_V = \mathbf{0}_V$.)

Onto and One-to-One Transformations

Recall that in Section 7 we expressed existence and uniqueness questions for matrix equations in terms of one-to-one and onto properties of matrix transformations. The question about the existence of a solution to the matrix equation $A\mathbf{x} = \mathbf{b}$ for any vector \mathbf{b} , where A is an $m \times n$ matrix, is also a question about the existence of a vector \mathbf{x} so that $T(\mathbf{x}) = \mathbf{b}$, where $T(\mathbf{x}) = A\mathbf{x}$. If, for each \mathbf{b} in \mathbb{R}^m there is at least one \mathbf{x} with $T(\mathbf{x}) = \mathbf{b}$, then T is an onto transformation. We can make a similar definition for any linear transformation.

Definition 35.2. A linear transformation T from a vector space V to a vector space W is **onto** if each \mathbf{b} in W is the image of at least one \mathbf{x} in V .

Similarly, the uniqueness of a solution to $A\mathbf{x} = \mathbf{b}$ for any \mathbf{b} in $\text{Col } A$ is the same as saying that for any \mathbf{b} in \mathbb{R}^m , there is at most one \mathbf{x} in \mathbb{R}^n such that $T(\mathbf{x}) = \mathbf{b}$. A matrix transformation with this property is one-to-one, and we can make a similar definition for any linear transformation.

Definition 35.3. A linear transformation T from a vector space V to a vector space W is **one-to-one** if each \mathbf{b} in W is the image of at most one \mathbf{x} in V .

With matrix transformations we saw that there are easy pivot criteria for determining whether a matrix transformation is one-to-one (a pivot in each column) or onto (a pivot in each row). If a linear transformation can be represented as a matrix transformation, then we can use these ideas. However, not every general linear transformation can be easily viewed as a matrix transformation, and in those cases we might have to resort to applying the definitions directly.

Activity 35.1. For each of the following transformations, determine if T is one-to-one and/or onto.

(a) $T : \mathbb{P}_2 \rightarrow \mathbb{P}_1$ defined by $T(a_0 + a_1t + a_2t^2) = a_0 + (a_1 + a_2)t$.

(b) $T : \mathbb{D} \rightarrow \mathcal{F}$ defined by $T(f) = f'$.

The Kernel and Range of Linear Transformation

As we saw in Preview Activity 35.1, any linear transformation sends the additive identity to the additive identity. If T is a matrix transformation defined by $T(\mathbf{x}) = A\mathbf{x}$ for some $m \times n$ matrix A , then we have seen that the set of vectors that T maps to the zero vector is $\text{Nul } A = \{\mathbf{x} : A\mathbf{x} = \mathbf{0}\}$, which is also $\text{Ker}(T) = \{\mathbf{x} : T(\mathbf{x}) = \mathbf{0}\}$. We can extend this idea of the kernel of a matrix transformation to any linear transformation T .

Definition 35.4. Let $T : V \rightarrow W$ be a linear transformation from the vector space V to the vector space W . The **kernel** of T is the set

$$\text{Ker}(T) = \{\mathbf{x} \in V : T(\mathbf{x}) = \mathbf{0}_W\},$$

where $\mathbf{0}_W$ is the additive identity in W .

Just as the null space of an $m \times n$ matrix A is a subspace of \mathbb{R}^n , the kernel of a linear transformation from a vector space V to a vector space W is a subspace of V . The proof is left to the exercises.

Theorem 35.5. Let $T : V \rightarrow W$ be a linear transformation from a vector space V to vector space W . Then $\text{Ker}(T)$ is a subspace of V .

The kernel of a linear transformation provides a convenient way to determine if the linear transformation is one-to-one. If T is one-to-one, then the only solution to $T(\mathbf{x}) = \mathbf{0}_W$ is $\mathbf{0}_V$ and $\text{Ker}(T)$ contains only the zero vector. If T is not one-to-one (and the domain of T is not just $\{\mathbf{0}\}$), then the number of solutions to $T(\mathbf{x}) = \mathbf{0}_W$ is infinite and $\text{Ker}(T)$ contains more than just the zero vector. We formalize this idea in the next theorem. (Compare to Theorem 13.3.) The formal proof is left for the exercises.

Theorem 35.6. A linear transformation T from a vector space V to a vector space W is one-to-one if and only if $\text{Ker}(T) = \{\mathbf{0}_V\}$, where $\mathbf{0}_V$ is the additive identity in V .

Activity 35.2.

- (a) Let $T : \mathbb{P}_1 \rightarrow \mathbb{P}_2$ be defined by $T(a_0 + a_1t) = a_1t^2$. Find $\text{Ker}(T)$. Is T one-to-one? Explain.

- (b) Let $T : \mathbb{D} \rightarrow \mathcal{F}$ be defined by $T(f) = f'$. Find $\text{Ker}(T)$.

Recall that the matrix-vector product $A\mathbf{x}$ is a linear combination of the columns of A and the set of all vectors of the form $A\mathbf{x}$ is the column space of A . For the matrix transformation T defined by $T(\mathbf{x}) = A\mathbf{x}$, the set of all vectors of the form $A\mathbf{x}$ is also the range of the transformation T . We can extend this idea to arbitrary linear transformations to define the *range* of a transformation.



Definition 35.7. Let $T : V \rightarrow W$ be a linear transformation from the vector space V to the vector space W . The **range** of T is the set

$$\text{Range}(T) = \{T(\mathbf{x}) : \mathbf{x} \text{ is in } V\}.$$

If $T(\mathbf{x}) = A\mathbf{x}$ for some $m \times n$ matrix A , we know that $\text{Col } A$, the span of the columns of A , is a subspace of \mathbb{R}^m . Consequently, the range of T , which is also the column space of A , is a subspace of \mathbb{R}^m . In general, the range of a linear transformation T from V to W is a subspace of W .

Theorem 35.8. Let $T : V \rightarrow W$ be a linear transformation from a vector space V to vector space W . Then $\text{Range}(T)$ is a subspace of W .

Proof. Let $T : V \rightarrow W$ be a linear transformation from a vector space V to vector space W . We have already shown that $T(\mathbf{0}_V) = \mathbf{0}_W$, so $\mathbf{0}_W$ is in $\text{Range}(T)$. To show that $\text{Range}(T)$ is a subspace of W we must also demonstrate that $\mathbf{w} + \mathbf{z}$ is in $\text{Range}(T)$ whenever \mathbf{w} and \mathbf{z} are in $\text{Range}(T)$ and that $a\mathbf{w}$ is in $\text{Range}(T)$ whenever a is a scalar and \mathbf{w} is in $\text{Range}(T)$. Let \mathbf{w} and \mathbf{z} be in $\text{Range}(T)$. Then $T(\mathbf{u}) = \mathbf{w}$ and $T(\mathbf{v}) = \mathbf{z}$ for some vectors \mathbf{u} and \mathbf{v} in V . Since T is a linear transformation, it follows that

$$T(\mathbf{u} + \mathbf{v}) = T(\mathbf{u}) + T(\mathbf{v}) = \mathbf{w} + \mathbf{z}.$$

So $\mathbf{w} + \mathbf{z}$ is in $\text{Range}(T)$.

Finally, let a be a scalar. The linearity of T gives us

$$T(a\mathbf{u}) = aT(\mathbf{u}) = a\mathbf{w},$$

so $a\mathbf{w}$ is in $\text{Range}(T)$. We conclude that $\text{Range}(T)$ is a subspace of W . ■

The subspace $\text{Range}(T)$ provides us with a convenient criterion for the transformation T being onto. The transformation T is onto if for each \mathbf{b} in W , there is at least one \mathbf{x} for which $T(\mathbf{x}) = \mathbf{b}$. This means that every \mathbf{b} in W belongs to $\text{Range}(T)$ for T to be onto.

Theorem 35.9. A linear transformation T from a vector space V to a vector space W is onto if and only if $\text{Range}(T) = W$.

Activity 35.3. Let $T : \mathbb{P}_1 \rightarrow \mathbb{P}_2$ be defined by $T(a_0 + a_1t) = a_1t^2$ as in Activity 35.2 . Describe the vectors in $\text{Range}(T)$. Is T onto? Explain.

Isomorphisms

If V is a vector space with a basis $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$, we have seen that the coordinate transformation $T : V \rightarrow \mathbb{R}^n$ defined by $T(\mathbf{x}) = [\mathbf{x}]_{\mathcal{B}}$ is a linear transformation that is both one-to-one and onto. This allows us to uniquely identify any vector in V with a vector in \mathbb{R}^n so that the vector space structure is preserved. In other words, the vector space V is for all intents and purposes the same as the vector space \mathbb{R}^n , except for the way we represent the vectors. This is a very powerful idea in that it shows that any vector space of dimension n is essentially \mathbb{R}^n and, consequently, any two vectors spaces of dimension n are essentially the same space. When this happens we say that the vectors spaces are *isomorphic* and call the coordinate transformation an *isomorphism*.



Definition 35.10. An **isomorphism** from a vector space V to a vector space W is a linear transformation $T : V \rightarrow W$ that is one-to-one and onto.

Activity 35.4. Assume that each of the following maps is a linear transformation. Which, if any, is an isomorphism? Justify your reasoning.

(a) $T : \mathbb{P}_1 \rightarrow \mathbb{R}^2$ defined by $T(a_0 + a_1 t) = \begin{bmatrix} a_1 \\ a_1 + a_0 \end{bmatrix}$

(b) $T : \mathcal{M}_{2 \times 2} \rightarrow \mathbb{P}_3$ defined by $T \left(\begin{bmatrix} a & b \\ c & d \end{bmatrix} \right) = a + bt + ct^2 + dt^3$.

(c) $T : \mathbb{R}^2 \rightarrow \mathbb{P}_2$ defined by $T \left(\begin{bmatrix} a \\ b \end{bmatrix} \right) = a + bt + at^2$.

It is left for the exercises to show that if $T : V \rightarrow W$ is an isomorphism, then $T^{-1} : W \rightarrow V$ is also an isomorphism. So if there is an isomorphism from a vector space V to a vector space W , we say that V and W are *isomorphic* vector spaces. It is also true that any vector space is isomorphic to itself, and that if V is isomorphic to W and W is isomorphic to U , then V and U are also isomorphic. The proof of this result is left to the exercises.

Examples

What follows are worked examples that use the concepts from this section.

Example 35.11. Let $T : \mathbb{P}_2 \rightarrow \mathbb{P}_3$ be defined by $T(p(t)) = tp(t) + p(0)$.

- (a) Show that T is a linear transformation.
- (b) Is T one-to-one? Justify your answer.
- (c)
 - i. Find three different polynomials in $\text{Range}(T)$.
 - ii. Find, if possible, a polynomial that is not in $\text{Range}(T)$.
 - iii. Describe $\text{Range}(T)$. What is $\dim(\text{Range}(T))$? Is T an onto transformation? Explain.

Example Solution.

- (a) To show that T is a linear transformation we must show that $T(p(t) + q(t)) = T(p(t)) + T(q(t))$ and $T(cp(t)) = cT(p(t))$ for every $p(t), q(t)$ in \mathbb{P}_2 and any scalar c . Let $p(t)$ and $q(t)$ be in \mathbb{P}_2 . Then

$$\begin{aligned} T(p(t) + q(t)) &= t(p(t) + q(t)) + (p(0) + q(0)) \\ &= (tp(t) + p(0)) + (tq(t) + q(0)) \\ &= T(p(t)) + T(q(t)) \end{aligned}$$

and

$$T(cp(t)) = tc p(t) + cp(0) = c(tp(t) + p(0)) = cT(p(t)).$$

Therefore, T is a linear transformation.



- (b) To determine if T is one-to-one, we find $\text{Ker}(T)$. Suppose $p(t) = a_0 + a_1t + a_2t^2 \in \text{Ker}(T)$. Then

$$0 = T(p(t)) = tp(t) + p(0) = a_0t + a_1t^2 + a_2t^3 + a_0.$$

Equating coefficients on like powers of t shows that $a_0 = a_1 = a_2 = 0$. Thus, $p(t) = 0$ and $\text{Ker}(T) = \{0\}$. Thus, T is one-to-one.

- (c) i. We can find three polynomials in $\text{Range}(T)$ by applying T to three different polynomials in \mathbb{P}_2 . So three polynomials in $\text{Range}(T)$ are

$$\begin{aligned} T(1) &= t + 1 \\ T(t) &= t^2 \\ T(t^2) &= t^3. \end{aligned}$$

- ii. A polynomial $q(t) = b_0 + b_1t + b_2t^2 + b_3t^3$ is in $\text{Range}(T)$ if $q(t) = T(p(t))$ for some polynomial $p(t) = a_0 + a_1t + a_2t^2$ in \mathbb{P}_2 . This would require that

$$b_0 + b_1t + b_2t^2 + b_3t^3 = a_0t + a_1t^2 + a_2t^3 + a_0.$$

But this would mean that $b_0 = a_0 = b_1$. So the polynomial $1 + 2t + t^2 + t^3$ is not in $\text{Range}(T)$.

- iii. Let $q(t)$ be in $\text{Range}(T)$. Then $q(t) = T(p(t))$ for some polynomial $p(t) = a_0 + a_1t + a_2t^2 \in \mathbb{P}_2$. Thus,

$$\begin{aligned} q(t) &= T(p(t)) \\ &= tp(t) + p(0) \\ &= a_0t + a_1t^2 + a_2t^3 + a_0 \\ &= a_0(t + 1) + a_1t^2 + a_2t^3. \end{aligned}$$

Therefore, $\text{Range}(T) = \text{Span}\{t + 1, t^2, t^3\}$. Since the reduced row echelon form

of $\left[\begin{array}{ccc} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{array} \right]$ is $\left[\begin{array}{ccc} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{array} \right]$, we conclude that the set $\{t + 1, t^2, t^3\}$ is linearly

independent. Thus, $\dim(\text{Range}(T)) = 3$. Since $\text{Range}(T)$ is a three-dimensional subspace of the four-dimensional space \mathbb{P}_3 , it follows that T is not onto.

Example 35.12. Let $T : \mathcal{M}_{2 \times 2} \rightarrow \mathbb{P}_3$ be defined by

$$T \left(\left[\begin{array}{cc} a & b \\ c & d \end{array} \right] \right) = a + bt + ct^2 + dt^3.$$

Show that T is an isomorphism from $\mathcal{M}_{2 \times 2}$ to \mathbb{P}_3 .

Example Solution.

We need to show that T is a linear transformation, and that T is both one-to-one and onto. We start by demonstrating that T is a linear transformation. Let $A = [a_{ij}]$ and $B = [b_{ij}]$ be in $\mathcal{M}_{2 \times 2}$.



Then

$$\begin{aligned}
 T(A + B) &= T([a_{ij} + b_{ij}]) \\
 &= (a_{11} + b_{11}) + (a_{12} + b_{12})t + (a_{21} + b_{21})t^2 + (a_{22} + b_{22})t^3 \\
 &= (a_{11} + a_{12}t + a_{21}t^2 + a_{22}t^3) + (b_{11} + b_{12}t + b_{21}t^2 + b_{22}t^3) \\
 &= T(A) + T(B).
 \end{aligned}$$

Now let c be a scalar. Then

$$\begin{aligned}
 T(cA) &= T([ca_{ij}]) \\
 &= (ca_{11}) + (ca_{12})t + (ca_{21})t^2 + (ca_{22})t^3 \\
 &= c(a_{11} + a_{12}t + a_{21}t^2 + a_{22}t^3) \\
 &= cT(A).
 \end{aligned}$$

Therefore, T is a linear transformation.

Next we determine $\text{Ker}(T)$. Suppose $A \in \text{Ker}(T)$. Then

$$0 = T(A) = a_{11} + a_{12}t + a_{21}t^2 + a_{22}t^3.$$

Equating coefficients of like powers of t shows that $a_{ij} = 0$ for each i and j . Thus, $A = 0$ and $\text{Ker}(T) = \{0\}$. It follows that T is one-to-one.

Finally, we show that $\text{Range}(T) = \mathbb{P}_3$. If $q(t) = a+bt+ct^2+dt^3 \in \mathbb{P}_3$, then $T\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) = q(t)$. So every polynomial in \mathbb{P}_3 is the image of some matrix in $\mathcal{M}_{2 \times 2}$. It follows that T is onto and also that T is an isomorphism.

Summary

Let T be a linear transformation from a vector space V to a vector space W .

- A function T from a vector space V to a vector space W is a linear transformation if
 - (1) $T(\mathbf{u} + \mathbf{v}) = T(\mathbf{u}) + T(\mathbf{v})$ for all \mathbf{u} and \mathbf{v} in V and
 - (2) $T(c\mathbf{u}) = cT(\mathbf{u})$ for all \mathbf{u} in V and all scalars c .
- Let T be a linear transformation from a vector space V to a vector space W . The kernel of T is the set

$$\text{Ker}(T) = \{\mathbf{x} \in V : T(\mathbf{x}) = \mathbf{0}\}.$$

The kernel of T is a subspace of V .

- Let T be a linear transformation from a vector space V to a vector space W . The transformation T is one-to-one if every vector in W is the image under T of at most one vector in V . A linear transformation T is one-to-one if and only if $\text{Ker}(T) = \{\mathbf{0}\}$.



- Let T be a linear transformation from a vector space V to a vector space W . The range of T is the set

$$\text{Range}(T) = \{T(\mathbf{v}) : \mathbf{v} \in V\}.$$

The range of T is a subspace of W .

- Let T be a linear transformation from a vector space V to a vector space W . The linear transformation T is onto if every vector in W is the image under T of at least one vector in V . The transformation T is onto if and only if $\text{Range}(T) = W$.
- An isomorphism from a vector space V to a vector space W is a linear transformation $T : V \rightarrow W$ that is one-to-one and onto.

Exercises

- We have seen that $C[a, b]$, the set of all continuous functions on the interval $[a, b]$ is a vector space. Let $T : C[a, b] \rightarrow \mathbb{R}$ be defined by $T(f) = \int_a^b f(x) dx$. Show that T is a linear transformation.
- If V is a vector space, prove that the identity map $I : V \rightarrow V$ defined by $I(\mathbf{x}) = \mathbf{x}$ for all $\mathbf{x} \in V$ is an isomorphism.
- Let V and W be vector spaces and let $T : V \rightarrow W$ be an isomorphism. Since T is one-to-one, it has an inverse function T^{-1} that has the property that $T^{-1}(\mathbf{w}) = \mathbf{v}$ whenever $T(\mathbf{v}) = \mathbf{w}$. Show that T^{-1} is also an isomorphism.
- Let U , V , and W be vector spaces, and let $S : U \rightarrow V$ and $T : V \rightarrow W$ be linear transformations. Determine which of the following is true. If a statement is true, verify it. If a statement is false, give a counterexample. (The result of this exercise, along with Exercises 2 and 3., shows that the isomorphism relation is reflexive, symmetric, and transitive. Thus, isomorphism is an equivalence relation.)
 - $T \circ S$ is a linear transformation
 - $S \circ T$ is a linear transformation
 - If S and T are one-to-one, then $T \circ S$ is one-to-one.
 - If S and T are onto, then $T \circ S$ is onto.
 - If S and T are isomorphisms, then $T \circ S$ is an isomorphism.
- For each of the following maps, determine which is a linear transformation. If a mapping is not a linear transformation, provide a specific example to show that a property of a linear transformation is not satisfied. If a mapping is a linear transformation, verify that fact and then determine if the mapping is one-to-one and/or onto. Throughout, let V and W be vector spaces.
 - $T : V \rightarrow V$ defined by $T(\mathbf{x}) = -\mathbf{x}$
 - $T : V \rightarrow V$ defined by $T(\mathbf{x}) = 2\mathbf{x}$



- (6) Let V be a vector space. In this exercise, we will investigate mappings $T : V \rightarrow V$ of the form $T(\mathbf{x}) = k\mathbf{x}$, where k is a scalar.
- For which values, if any, of k is T a linear transformation?
 - For which values of k , if any, is T an isomorphism?
- (7) Let V be a finite-dimensional vector space and W a vector space. Show that if $T : V \rightarrow W$ is an isomorphism, and $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$ is a basis for V , then $\mathcal{C} = \{T(\mathbf{v}_1), T(\mathbf{v}_2), \dots, T(\mathbf{v}_n)\}$ is a basis for W . Hence, if V is a vector space of dimension n and W is isomorphic to V , then $\dim(W) = n$ as well.
- (8) Let $W = \left\{ \begin{bmatrix} x & -x \\ -x & x \end{bmatrix} : x \text{ is a scalar} \right\}$.
- Show that W is a subspace of $\mathcal{M}_{2 \times 2}$.
 - The space W is isomorphic to \mathbb{R}^k for some k . Determine the value of k and explain why W is isomorphic to \mathbb{R}^k .
- (9) Is it possible for a vector space to be isomorphic to one of its proper subspaces? Justify your answer.
- (10) Prove Theorem 35.5 (Hint: See Preview Activity 13.1.)
- (11) Prove Theorem 35.6.
- (12) Let V and W be finite dimensional vector spaces, and let $T : V \rightarrow W$ be a linear transformation. Show that T is uniquely determined by its action on a basis for V . That is, if $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$ is a basis for V , and we know $T(\mathbf{v}_1), T(\mathbf{v}_2), \dots, T(\mathbf{v}_n)$, then we know exactly which vector $T(\mathbf{u})$ is for any \mathbf{u} in V .
- (13) Let V and W be vectors spaces and let $T : V \rightarrow W$ be a linear transformation. If V' is a subspace of V , let
- $$T(V') = \{T(\mathbf{v}) : \mathbf{v} \text{ is in } V'\}.$$
- Show that $T(V')$ is a subspace of W .
 - Suppose $\dim(V') = n$. Show that $\dim(T(V')) \leq n$. Can $\dim(T(V')) = n$ even if T is not an isomorphism? Justify your answer.
- (14) Let V and W be vector spaces and define $L(V, W)$ to be the set of all linear transformations from V to W . If S and T are in $L(V, W)$ and c is a scalar, then define $S + T$ and cT as follows:
- $$(S + T)(\mathbf{v}) = S(\mathbf{v}) + T(\mathbf{v}) \quad \text{and} \quad (cT)(\mathbf{v}) = cT(\mathbf{v})$$
- for all \mathbf{v} in V . Prove that $L(V, W)$ is a vector space.
- (15) The Rank-Nullity Theorem (Theorem 15.5) states that the rank plus the nullity of a matrix equals the number of columns of the matrix. There is a similar result for linear transformations that we prove in this exercise. Let $T : V \rightarrow W$ be a linear transformation from an n dimensional vector space V to an m dimensional vector space W .



Show that $\dim(\text{Ker}(T)) + \dim(\text{Range}(T)) = n$. (Hint: Let $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k\}$ be a basis for $\text{Ker}(T)$. Extend this basis to a basis $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k, \mathbf{v}_{k+1}, \dots, \mathbf{v}_n\}$ of V . Use this basis to find a basis for $\text{Range}(T)$.)

- (16) Let V and W be vector spaces with $\dim(V) = n = \dim(W)$, and let $T : V \rightarrow W$ be a linear transformation
- Prove or disprove: If T is one-to-one, then T is also onto. (Hint: Use Exercise 3.)
 - Prove or disprove: If T is onto, then T is also one-to-one. (Hint: Use Exercise 3.)
- (17) Suppose \mathcal{B}_1 is a basis of an n -dimensional vector space V , and \mathcal{B}_2 is a basis of an n -dimensional vector space W . Show that the map T_{VW} which sends every \mathbf{x} in V to the vector \mathbf{y} in W such that $[\mathbf{x}]_{\mathcal{B}_1} = [\mathbf{y}]_{\mathcal{B}_2}$ is a linear transformation. (Combined with Exercise 16 in Section 25, we can conclude that T_{VW} is an isomorphism. Thus, any two vectors space of dimension n are isomorphic.)
- (18) Label each of the following statements as True or False. Provide justification for your response.
- True/False** The mapping $T : \mathcal{M}_{2 \times 2} \rightarrow \mathbb{R}$ defined by $T(A) = \det(A)$ is a linear transformation.
 - True/False** The mapping $T : \mathcal{M}_{2 \times 2} \rightarrow \mathcal{M}_{2 \times 2}$ defined by $T(A) = A^{-1}$ if A is invertible and $T(A) = 0$ (the zero matrix) otherwise is a linear transformation.
 - True/False** Let V and W be vector spaces. The mapping $T : V \rightarrow W$ defined by $T(\mathbf{x}) = \mathbf{0}$ is a linear transformation.
 - True/False** If $T : V \rightarrow W$ is a linear transformation, and if $\{\mathbf{v}_1, \mathbf{v}_2\}$ is a linearly independent set in V , then the set $\{T(\mathbf{v}_1), T(\mathbf{v}_2)\}$ is a linearly independent set in W .
 - True/False** A one-to-one transformation is a transformation where each input has a unique output.
 - True/False** A one-to-one linear transformation is a transformation where each output can only come from a unique input.
 - True/False** Let V and W be vector spaces with $\dim(V) = 2$ and $\dim(W) = 3$. A linear transformation from V to W cannot be onto.
 - True/False** Let V and W be vector spaces with $\dim(V) = 3$ and $\dim(W) = 2$. A linear transformation from V to W cannot be onto.
 - True/False** Let V and W be vector spaces with $\dim(V) = 3$ and $\dim(W) = 2$. A linear transformation from V to W cannot be one-to-one.
 - True/False** If \mathbf{w} is in the range of a linear transformation T , then there is an \mathbf{x} in the domain of T such that $T(\mathbf{x}) = \mathbf{w}$.
 - True/False** If T is a one-to-one linear transformation, then $T(\mathbf{x}) = \mathbf{0}$ has a non-trivial solution.

Project: Fractals via Iterated Function Systems

In this project we will see how to use linear transformations to create iterated functions systems to generate fractals. We illustrate the idea with the Sierpinski triangle, an approximate picture of which is shown at left in Figure 35.2. To make this figure, we need to identify linear transformations

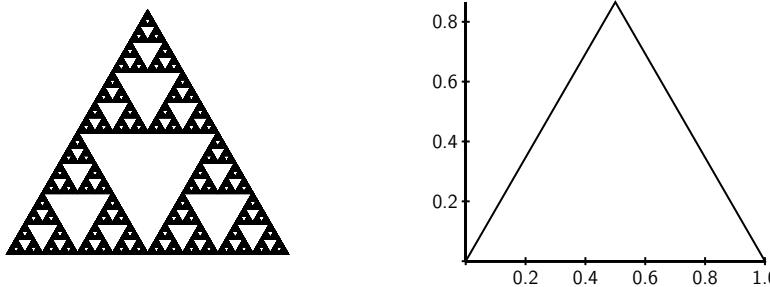


Figure 35.2: Left: An approximation of the Sierpinski triangle. Right: A triangle.

that can be put together to produce the Sierpinski triangle.

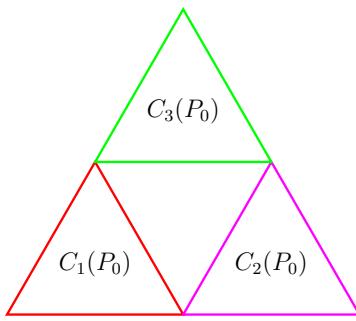
Project Activity 35.1. Let $\mathbf{v}_1 = [0 \ 0]^\top$, $\mathbf{v}_2 = [1 \ 0]^\top$, and $\mathbf{v}_3 = \left[\frac{1}{2} \ \frac{\sqrt{3}}{4}\right]^\top$ be three vectors in the plane whose endpoints form the vertices of a triangle P_0 as shown at right in Figure 35.2. Let $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ be the linear transformation defined by $T(\mathbf{x}) = A\mathbf{x}$, where $A = \frac{1}{2}I_2 = \frac{1}{2} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$.

- What are $T(\mathbf{v}_1)$, $T(\mathbf{v}_2)$, and $T(\mathbf{v}_3)$? Draw a picture of the figure $T(P_0)$ whose vertices are the endpoint of these three vectors. How is $T(P_0)$ related to P_0 ?
- Since the transformation T shrinks objects, we call T (or A) a *contraction mapping*. Notice that when we apply T to P_0 it creates a smaller version of P_0 . The next step is to use T to make three smaller copies of P_0 , and then translate these copies to the vertices of P_0 . A translation can be performed by adding a vector to $T(P_0)$. Find C_1 so that $C_1(P_0)$ is a copy of P_0 half the size and translated so that $C_1(\mathbf{v}_1) = \mathbf{v}_1$. Then find C_2 so that $C_2(P_0)$ is a copy of P_0 half the size and translated so that $C_2(\mathbf{v}_2) = \mathbf{v}_2$. Finally, find C_3 so that $C_3(P_0)$ is a copy of P_0 half the size and translated so that $C_3(\mathbf{v}_3) = \mathbf{v}_3$. Draw pictures of each to illustrate.

Project Activity 35.1 contains the information we need to create an iterated function system to produce the Sierpinski triangle. One more step should help us understand the general process.

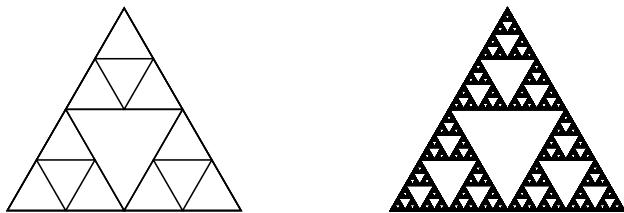
Project Activity 35.2. Using the results of Project Activity 35.1, define $P_{1,i}$ to be $C_i(P_0)$ for each i . That is, $P_{1,1} = C_1(P_0)$, $P_{1,2} = C_2(P_0)$, and $P_{1,3} = C_3(P_0)$. So $P_{1,i}$ is a triangle half the size of the original translated to the i th vertex of the original. Let $P_1 = \bigcup_{i=1}^3 P_{1,i}$. That is, P_1 is the union of the shaded triangles in Figure 35.3.

- Apply C_1 from Project Activity 35.1 to P_1 . What is the resulting figure? Draw a picture to illustrate.

Figure 35.3: $C_1(P_0)$, $C_2(P_0)$, and $C_3(P_0)$.

- (b) Apply C_2 from Project Activity 35.1 to P_1 . What is the resulting figure? Draw a picture to illustrate.
- (c) Apply C_3 from Project Activity 35.1 to P_1 . What is the resulting figure? Draw a picture to illustrate.

The procedures from Activities 35.1 and 35.2 can be continued, replacing P_0 with P_1 , then P_2 , and so on. In other words, for $i = 1, 2$, and 3 , let $P_{2,i} = C_i(P_1)$. Then let $P_2 = \bigcup_{i=1}^3 P_{2,i}$. A picture of P_2 is shown at left in Figure 35.4. We can continue this procedure, each time replacing P_{j-1} with P_j . A picture of P_9 is shown at right in Figure 35.4.

Figure 35.4: Left: P_2 . Right: P_9 .

If we continue this process, taking the limit as i approaches infinity, the resulting sequence of sets converges to a fixed set, in this case the famous Sierpinski triangle. So the picture of P_9 in Figure 35.4 is a close approximation of the Sierpinski triangle. This algorithm for building the Sierpinski triangle is called the *deterministic algorithm*. A Sage cell to illustrate this algorithm for producing approximations to the Sierpinski triangle can be found at http://faculty.gvsu.edu/schlicks/STriangle_Sage.html.

In general, an iterated function system (IFS) is a finite set $\{f_1, f_2, \dots, f_m\}$ of contraction mappings from \mathbb{R}^2 to \mathbb{R}^2 . If we start with a set S_0 , and let $S_1 = \bigcup f_i(S_0)$, $S_2 = \bigcup f_i(S_1)$, and so on, then in “nice” cases (we won’t concern ourselves with what “nice” means here), the sequence S_0, S_1, S_2, \dots converges in some sense to a fixed set. That fixed set is called the *attractor* of the iterated function system. It is that attractor that is the fractal. It is fascinating to note that our starting set does not matter, the same attractor is obtained no matter which set we use as S_0 .

One aspect of a fractal is that fractals are self-similar, that is they are made up of constituent pieces that look just like the whole. More specifically, a subset S of \mathbb{R}^2 is *self-similar* if it can be expressed in the form

$$S = S_1 \cup S_2 \cup \cdots \cup S_k$$

for non-overlapping sets S_1, S_2, \dots, S_k , each of which is congruent to S by the same scaling factor. So, for example, the Sierpinski triangle is self-similar, made up of three copies of the whole, each contracted by a factor of 2.

If we have an IFS, then we can determine the attractor by drawing the sequence of sets that the IFS generates. A more interesting problem is, given a self-similar figure, whether we can construct an IFS that has that figure as its attractor.

Project Activity 35.3. A picture of an emerging Sierpinski carpet is shown at left in Figure 35.5. A Sage cell to illustrate this algorithm for producing approximations to the Sierpinski carpet can be found at http://faculty.gvsu.edu/schlicks/SCarpet_Sage.html. In this activity we will see how to find an iterated function system that will generate this fractal.

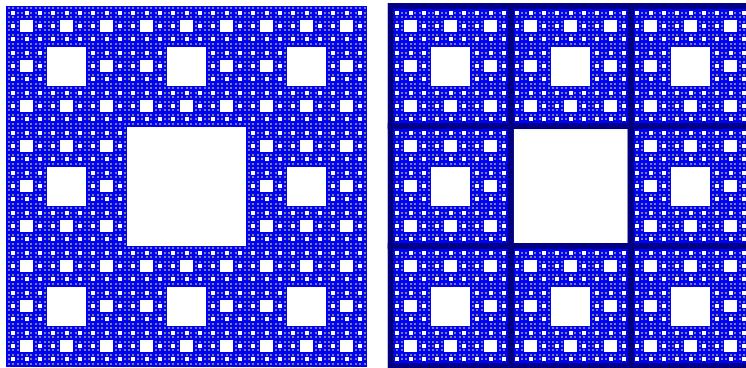


Figure 35.5: A Sierpinski carpet.

- (a) To create an IFS to generate this fractal, we need to understand how many self-similar pieces make up this figure. Use the image at right in Figure 35.5 to determine how many pieces we need.
- (b) For each of the self-similar pieces identified in part (a), find a linear transformation and a translation that maps the entire figure to the self-similar piece. (Hint: You could assume that the carpet is embedded in the unit square.)
- (c) Test your IFS to make sure that it actually generates the Sierpinski carpet. There are many websites that allow you to do this, one of which is <http://cs.lmu.edu/~ray/notes/ifs/>. In this program, the mapping f defined by

$$f(\mathbf{x}) = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \mathbf{x} + \begin{bmatrix} e \\ f \end{bmatrix}$$

is represented as the string

$$a \ b \ c \ d \ e \ f.$$

Most programs generally use a different algorithm to create the attractor, plotting points instead of sets. In this algorithm, each contraction mapping is assigned a probability as well (the larger parts of the figure are usually given higher probabilities), so you will enter each contraction mapping in the form

$$a \ b \ c \ d \ e \ f \ p$$

where p is the probability attached to that contraction mapping. As an example, the IFS code for the Sierpinski triangle is

$$\begin{array}{ccccccc} 0.5 & 0 & 0 & 0.5 & 0 & 0 & 0.33 \\ 0.5 & 0 & 0 & 0.5 & 1 & 0 & 0.33 \\ 0.5 & 0 & 0 & 0.5 & 0.5 & 0.43 & 0.33 \end{array}$$

The contraction mappings we have used so far only involve contractions and translations. But we do not have to restrict ourselves to just these types of contraction mappings.

Project Activity 35.4. Consider the fractal represented in Figure 35.6. Find an IFS that generates this fractal. Explain your reasoning. (Hint: Two reflections are involved.) Check with a fractal generator to ensure that you have an appropriate IFS.

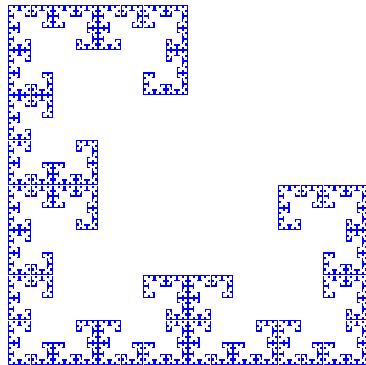


Figure 35.6: A fractal.

We conclude our construction of fractals with one more example. The contraction mappings in iterated function can also involve rotations.

Project Activity 35.5. Consider the Lévy Dragon fractal shown at left in Figure 35.7. Find an IFS that generates this fractal. Explain your reasoning. (Hint: Two rotations are involved – think of the fractal as contained in a blue triangle as shown at right in Figure 35.7.) Check with a fractal generator to ensure that you have an appropriate IFS.

We end this project with a short discussion of fractal dimension. The fractals we have seen are very strange in many ways, one of which is dimension. We have studied the dimensions of subspaces of R^n – each subspace has an integer dimension that is determined by the number of elements in a basis. Fractals also have a dimension, but the dimension of a fractal is generally not

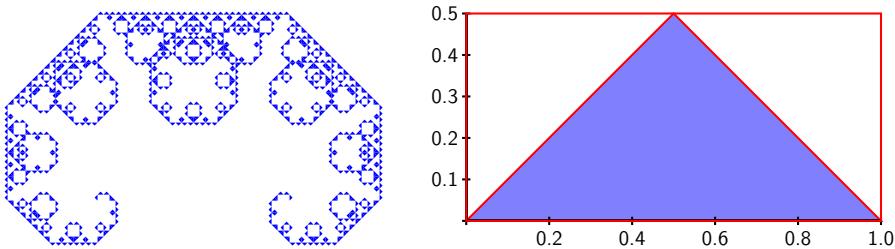


Figure 35.7: The Lévy Dragon.

an integer. To try to understand fractal dimension, notice that a line segment is self-similar. We can break up a line segment into 2 non-overlapping line segments of the same length, or 3, or 4, or any number of non-overlapping segments of the same length. A square is slightly different. We can partition a square into 4 non-overlapping squares, or 9 non-overlapping squares, or n^2 non-overlapping squares for any positive integer n as shown in Figure 35.8. Similarly, we can break

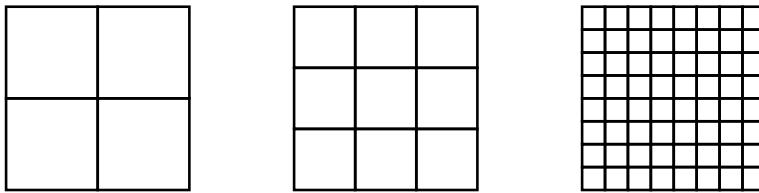


Figure 35.8: Self-similar squares.

up a cube into n^3 non-overlapping congruent cubes. A line segment lies in a one-dimensional space, a square in a two-dimensional space, and a cube in the three-dimensional space. Notice that these dimensions correspond to the exponent of the number of self-similar pieces with scaling n into which we can partition the object. We can use this idea of dimension in a way that we can apply to fractals. Let $d(\text{object})$ be the dimension of the object. We can partition a square into n^2 non-overlapping squares, so

$$d(\text{square}) = 2 = 2 \frac{\ln(n)}{\ln(n)} = \frac{\ln(n^2)}{\ln(n)} = \frac{\ln(\text{number of self-similar pieces})}{\ln(\text{scaling factor})}.$$

Similarly, for the cube we have

$$d(\text{cube}) = 3 = 3 \frac{\ln(n)}{\ln(n)} = \frac{\ln(n^3)}{\ln(n)} = \frac{\ln(\text{number of self-similar pieces})}{\ln(\text{scaling factor})}.$$

We can then take this as our definition of the dimension of a self-similar object, when the scaling factors are all the same (the fern fractal in Figure 35.1 is an example of a fractal generated by an iterated function system in which the scaling factors are not all the same).

Definition 35.13. The **fractal** or **Hausdorff** dimension h of a self-similar set S is

$$h(S) = \frac{\ln(\text{number of self-similar pieces})}{\ln(\text{scaling factor})}.$$

Project Activity 35.6. Find the fractal dimensions of the Sierpinski triangle and the Sierpinski carpet. These are well-known and you can look them up to check your result. Then find the fractal dimension of the fractal with IFS

$$\begin{array}{ccccccc} 0.38 & 0 & 0 & 0.38 & -0.59 & 0.81 & 0.2 \\ 0.38 & 0 & 0 & 0.38 & -0.95 & -0.31 & 0.2 \\ 0.38 & 0 & 0 & 0.38 & 0 & -1 & 0.2 \\ 0.38 & 0 & 0 & 0.38 & 0.95 & -0.31 & 0.2 \\ 0.38 & 0 & 0 & 0.38 & 0.59 & 0.81 & 0.2 \end{array} .$$

You might want to draw this fractal using an online generator.

Section 36

The Matrix of a Linear Transformation

Focus Questions

By the end of this section, you should be able to give precise and thorough answers to the questions listed below. You may want to keep these questions in mind to focus your thoughts as you complete the section.

- How do we represent a linear transformation from \mathbb{R}^n to \mathbb{R}^m as a matrix transformation?
- How can we represent any linear transformation from a finite dimensional vector space V to a finite dimensional vector space W as a matrix transformation?
- In what ways is representing a linear transformation as a matrix transformation useful?

Application: Secret Sharing Algorithms

Suppose we have a secret that we want or need to share with several individuals. For example, a bank manager might want to share pieces of the code for the bank vault with several employees so that if the manager is not available, some subset of these employees could open the vault if they work together. This allows for a significant amount of security while still making the code available if needed. As another example, in order to keep passwords secure (as they can be hard to remember), a person could implement a secret sharing scheme. The person would generate a set of shares for a given password and store them in several different places. If the person forgets the password, it can be reconstructed by collecting some set of these shares. Since the shares can be stored in many different places, the password is relatively secure.

The idea behind secret sharing algorithms is to break a secret into a number (n) of pieces and give each person one piece. A code is then created in such a way that any k individuals could combine their information and learn the secret, but no group of fewer than k individuals could. This is called a (k, n) threshold scheme. One secret sharing algorithm is Shamir's Secret Sharing,

which, as we will see later in this section, involves Lagrange polynomials. In order to implement this algorithm, we will use matrices of linear transformations from polynomial spaces to \mathbb{R}^n .

Introduction

A matrix transformation T from \mathbb{R}^n to \mathbb{R}^m is a linear transformation defined by $T(\mathbf{x}) = A\mathbf{x}$ for some $m \times n$ matrix A . We can use the matrix to quickly determine if T is one-to-one or onto – if every column of A contains a pivot, then T is one-to-one, and if every row of A contains a pivot, then T is onto. As we will see, we can represent any linear transformation between finite dimensional vector spaces as a matrix transformation, which will allow us to use all of the tools we have developed for matrices to study linear transformations. We will begin with linear transformations from \mathbb{R}^n to \mathbb{R}^m .

Preview Activity 36.1. Let T be a linear transformation from \mathbb{R}^2 into \mathbb{R}^3 . Let's also say that we know the following information about T :

$$T \left(\begin{bmatrix} 1 \\ 0 \end{bmatrix} \right) = \begin{bmatrix} 0 \\ 2 \\ -1 \end{bmatrix} \text{ and } T \left(\begin{bmatrix} 0 \\ 1 \end{bmatrix} \right) = \begin{bmatrix} 1 \\ 3 \\ 2 \end{bmatrix}.$$

(1) Find $T \left(\begin{bmatrix} 2 \\ 0 \end{bmatrix} \right)$. Hint: Use the fact that T is a linear transformation.

(2) Find $T \left(\begin{bmatrix} 1 \\ 1 \end{bmatrix} \right)$.

(3) Find $T \left(\begin{bmatrix} 2 \\ 3 \end{bmatrix} \right)$.

(4) Find $T \left(\begin{bmatrix} a \\ b \end{bmatrix} \right)$ for any real numbers a and b .

(5) Is it possible to find a matrix A so that $T \left(\begin{bmatrix} a \\ b \end{bmatrix} \right) = A \begin{bmatrix} a \\ b \end{bmatrix}$ for any real numbers a and b ? If so, what is this matrix? If not, why not?

Linear Transformations from \mathbb{R}^n to \mathbb{R}^m

Preview Activity 36.1 illustrates the method for representing a linear transformation from \mathbb{R}^n to \mathbb{R}^m as a matrix transformation. We now consider the general context.

Activity 36.1. Let T be a linear transformation from \mathbb{R}^n to \mathbb{R}^m . Let \mathbf{e}_i be the i th column of the



$n \times n$ identity matrix. In other words,

$$\mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \mathbf{e}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \dots, \mathbf{e}_n = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}$$

in \mathbb{R}^n . The set $\{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$ is the standard basis for \mathbb{R}^n . Let $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$ in \mathbb{R}^n . Explain why $T(\mathbf{x}) = A\mathbf{x}$ for every \mathbf{x} in \mathbb{R}^n , where

$$A = [T(\mathbf{e}_1) \ T(\mathbf{e}_2) \ \cdots \ T(\mathbf{e}_n)]$$

is the matrix with columns $T(\mathbf{e}_1), T(\mathbf{e}_2), \dots, T(\mathbf{e}_n)$.

As we discovered in Activity 36.1, a linear transformation from \mathbb{R}^n to \mathbb{R}^m can be expressed as a matrix transformation where the columns of the matrix are given by the images of \mathbf{e}_i , the columns of the $n \times n$ identity matrix. This matrix is called the *standard matrix* of the linear transformation.

Definition 36.1. If T is a linear transformation from \mathbb{R}^n to \mathbb{R}^m , the **standard matrix** for T is the matrix

$$A = [T(\mathbf{e}_1) \ T(\mathbf{e}_2) \ \cdots \ T(\mathbf{e}_n)]$$

with columns $T(\mathbf{e}_1), T(\mathbf{e}_2), \dots, T(\mathbf{e}_n)$, where $\{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$ is the standard basis for \mathbb{R}^n . More specifically, $T(\mathbf{x}) = A\mathbf{x}$ for every \mathbf{x} in \mathbb{R}^n .

The Matrix of a Linear Transformation

We saw in the previous section that any linear transformation T from \mathbb{R}^n to \mathbb{R}^m is in fact a matrix transformation. In this section, we turn to the general question – can any linear transformation $T : V \rightarrow W$ between an n -dimensional vector space V and an m -dimensional vector space W be represented in some way as a matrix transformation? This will allow us to extend ideas like eigenvalues and eigenvectors to arbitrary linear transformations. We begin by investigating an example.

The general process for representing a linear transformation as a matrix transformation is as described in Activity 36.2.

Activity 36.2. Let V be an n dimensional vector space and W an m dimensional vector space and suppose $T : V \rightarrow W$ is a linear transformation. Let $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$ be a basis for V and $\mathcal{C} = \{\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_m\}$ a basis for W . Let \mathbf{v} be in V so that

$$\mathbf{v} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_n\mathbf{v}_n$$

for some scalars c_1, c_2, \dots, c_n .



- (a) What is $[\mathbf{v}]_{\mathcal{B}}$?
- (b) Note that for \mathbf{v} in V , $T(\mathbf{v})$ is an element in W , so we can consider the coordinate vector for $T(\mathbf{v})$ with respect to \mathcal{C} . Explain why

$$[T(\mathbf{v})]_{\mathcal{C}} = c_1[T(\mathbf{v}_1)]_{\mathcal{C}} + c_2[T(\mathbf{v}_2)]_{\mathcal{C}} + \cdots + c_n[T(\mathbf{v}_n)]_{\mathcal{C}}. \quad (36.1)$$

- (c) Now find a matrix A so that Equation (36.1) has the form $[T(\mathbf{v})]_{\mathcal{C}} = A[\mathbf{v}]_{\mathcal{B}}$. We call the matrix A to be the *matrix for T relative to the bases \mathcal{B} and \mathcal{C}* and denote this matrix as $[T]_{\mathcal{B}}^{\mathcal{C}}$ so that

$$[T(\mathbf{v})]_{\mathcal{C}} = [T]_{\mathcal{B}}^{\mathcal{C}}[\mathbf{v}]_{\mathcal{B}}.$$

Activity 36.2 shows us how to represent a linear transformation as a matrix transformation.

Definition 36.2. Let T be a linear transformation from a vector space V with basis $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$ to a vector space W with basis $\mathcal{C} = \{\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_m\}$. The **matrix for T with respect to the bases \mathcal{B} and \mathcal{C}** is the matrix

$$[T]_{\mathcal{B}}^{\mathcal{C}} = [[T(\mathbf{v}_1)]_{\mathcal{C}} \ [T(\mathbf{v}_2)]_{\mathcal{C}} \ [T(\mathbf{v}_3)]_{\mathcal{C}} \ \cdots \ [T(\mathbf{v}_n)]_{\mathcal{C}}].$$

If $V = W$ and $\mathcal{C} = \mathcal{B}$, then we use the notation $[T]_{\mathcal{B}}$ as a shorthand for $[T]_{\mathcal{B}}^{\mathcal{C}}$. The matrix $[T]_{\mathcal{B}}^{\mathcal{C}}$ has the property that

$$[T(\mathbf{v})]_{\mathcal{C}} = A[\mathbf{v}]_{\mathcal{B}}$$

for any \mathbf{v} in V . Recall that we can find the unique vector in W whose coordinate vector is $[T(\mathbf{v})]_{\mathcal{C}}$, so we have completely realized the transformation T as a matrix transformation. In essence, we are viewing T as a composite of linear transformations, first from V to \mathbb{R}^n , then from \mathbb{R}^n to \mathbb{R}^m , then from \mathbb{R}^m to W as illustrated in Figure 36.1.

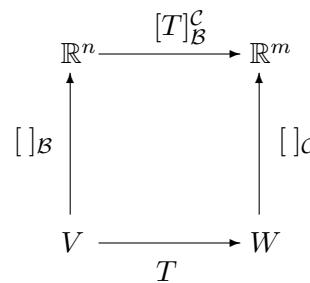


Figure 36.1: Visualizing the matrix of a linear transformation.

Activity 36.3. Let $T : \mathbb{P}_2 \rightarrow \mathbb{P}_1$ be defined by $T(p(t)) = p'(t)$. Let $\mathcal{B} = \{1 + t, 1 - t, t + t^2\}$ be a basis for \mathbb{P}_2 and $\mathcal{C} = \{1, t\}$ be a basis for \mathbb{P}_1 .

- (a) Find the matrix $[T]_{\mathcal{B}}^{\mathcal{C}}$.
- (b) Find $[1 + t + t^2]_{\mathcal{B}}$. Then use the matrix $[T]_{\mathcal{B}}^{\mathcal{C}}$ to calculate $[T(1 + t + t^2)]_{\mathcal{C}}$. Use the fact that $1 + t + t^2 = (1) + (t + t^2)$.

(c) Calculate $T(1 + t + t^2)$ directly from the definition of T and compare to your result from part (b).

(d) Find the matrix $[T]_{\mathcal{B}}^{\mathcal{C}'}$ where $\mathcal{C}' = \{1 + t, 1 - t\}$. Use the facts that

$$-1 = -\frac{1}{2}(1+t) - \frac{1}{2}(1-t) \text{ and } 1+2t = \frac{3}{2}(1+t) - \frac{1}{2}(1-t).$$

A Connection between $\text{Ker}(T)$ and a Matrix Representation of T

Recall that we defined the kernel of a matrix transformation $T : \mathbb{R}^n \rightarrow \mathbb{R}^m$ to be the set of vectors that T maps to the zero vector. If T is the matrix transformation defined by $T(\mathbf{x}) = A\mathbf{x}$, we also saw that the kernel of T is the same as the null space of A . We can make a similar connection between a linear transformation and a matrix that defines the transformation.

Activity 36.4. Let $T : V \rightarrow W$ be a linear transformation from an n -dimensional vector space V to an m -dimensional vector space W with additive identity $\mathbf{0}_W$. Let \mathcal{B} be a basis for V and \mathcal{C} a basis for W . Let $A = [T]_{\mathcal{B}}^{\mathcal{C}}$. Let T' be the matrix transformation from \mathbb{R}^n to \mathbb{R}^m defined by $T'(\mathbf{x}) = A\mathbf{x}$.

- (a) Show that if \mathbf{v} is in $\text{Ker}(T)$, then $[\mathbf{v}]_{\mathcal{B}}$ is in $\text{Nul } A$. (Hint: Apply an appropriate coordinate transformation.)
- (b) Let $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$ and $\mathcal{C} = \{\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_m\}$. Show that if the vector $\mathbf{x} = [x_1 \ x_2 \ \cdots \ x_n]$ is in $\text{Nul } A$, then the vector $\mathbf{v} = x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + \cdots + x_n\mathbf{v}_n$ is in $\text{Ker}(T)$. (Hint: Note that $[\mathbf{v}]_{\mathcal{B}} = \mathbf{x}$.)

Activity 36.4 shows that if $T : V \rightarrow W$ is a linear transformation from an n -dimensional vector space V with basis \mathcal{B} to an m -dimensional vector space W with basis \mathcal{C} , then there is a one-to-one correspondence between vectors in $\text{Ker}(T)$ and vectors in $\text{Nul } [T]_{\mathcal{B}}^{\mathcal{C}}$. Recall that T is one-to-one if and only if $\text{Ker}(T) = \{\mathbf{0}_V\}$, where $\mathbf{0}_V$ is the additive identity of V . So T will be one-to-one if and only if $\text{Nul } [T]_{\mathcal{B}}^{\mathcal{C}} = \{\mathbf{0}\}$. In other words, T is one-to-one if and only if every column of $\text{Nul } [T]_{\mathcal{B}}^{\mathcal{C}}$ is a pivot column. Note that this does not depend on the basis \mathcal{B} and \mathcal{C} .

A similar argument shows that T is onto if and only if every row of $\text{Nul } [T]_{\mathcal{B}}^{\mathcal{C}}$ contains a pivot. This is left to the exercises (see Exercise 7.).

Examples

What follows are worked examples that use the concepts from this section.

Example 36.3. Let $T : \mathbb{P}_1 \rightarrow \mathbb{P}_2$ be defined by $T(a + bt) = b + at + 2at^2$.

- (a) Show that T is a linear transformation.
- (b) Let $\mathcal{S}_1 = \{1, t\}$ and $\mathcal{S}_2 = \{1, t, t^2\}$ be the standard bases for \mathbb{P}_1 and \mathbb{P}_2 , respectively. Find the matrix $[T]_{\mathcal{S}_2}^{\mathcal{S}_1}$ for T with respect to \mathcal{S}_2 and \mathcal{S}_1 . Use the matrix $[T]_{\mathcal{S}_2}^{\mathcal{S}_1}$ to determine if T is one-to-one and/or onto. Explain your reasoning. Use technology as appropriate.



- (c) Use the matrix $[T]_{\mathcal{S}_2}^{\mathcal{S}_1}$ to find $\text{Ker}(T)$ and $\text{Range}(T)$. If possible, find a basis for each.
- (d) Let $\mathcal{B} = \{1 + t, 1 - t\}$ be a basis for \mathbb{P}_1 and $\mathcal{C} = \{t, 1 - t^2, t + t^2\}$ a basis for \mathbb{P}_2 . Find the matrix $[T]_{\mathcal{B}}^{\mathcal{C}}$ for T with respect to \mathcal{C} and \mathcal{B} .
- (e) Find $T(1 - 2t)$ using the matrix $[T]_{\mathcal{B}}^{\mathcal{C}}$.

Example Solution.

- (a) Let $p(t) = a + bt$ and $q(t) = c + dt$ be in \mathbb{P}_1 . Then

$$\begin{aligned} T(p(t) + q(t)) &= T((a + c) + (b + d)t) \\ &= (b + d) + (a + c)t + 2(a + c)t^2 \\ &= (b + at + 2at^2) + (d + ct + 2ct^2) \\ &= T(p(t)) + T(q(t)). \end{aligned}$$

Now let k be a scalar. Then

$$\begin{aligned} T(kp(t)) &= T((ka) + (kb)t) \\ &= (kb) + (ka)t + (2ka)t^2 \\ &= k(b + at + 2at^2) \\ &= kT(p(t)). \end{aligned}$$

We conclude that T is a linear transformation.

- (b) Recall that

$$[T]_{\mathcal{S}_1}^{\mathcal{S}_2} = [[T(1)]_{\mathcal{S}_2} [T(t)]_{\mathcal{S}_2}] = [[t + 2t^2]_{\mathcal{S}_2} [1]_{\mathcal{S}_2}] = \begin{bmatrix} 0 & 1 \\ 1 & 0 \\ 2 & 0 \end{bmatrix}.$$

The linear transformation T is one-to-one and/or onto if and only if the matrix transformation defined by $[T]_{\mathcal{S}_1}^{\mathcal{S}_2}$ is one-to-one and/or onto. The reduced row echelon form of $[T]_{\mathcal{S}_1}^{\mathcal{S}_2}$ is $\begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}$. Since $[T]_{\mathcal{S}_1}^{\mathcal{S}_2}$ has a pivot in every column, T is one-to-one. However, $[T]_{\mathcal{S}_1}^{\mathcal{S}_2}$ does not have a pivot in every row, so T is not onto.

- (c) From part (b) we know that T is one-to-one, so $\text{Ker}(T) = \{0\}$. Let $q(t)$ be a polynomial in $\text{Range}(T)$. Then there is a polynomial $p(t)$ in \mathbb{P}_1 such that $T(p(t)) = q(t)$. It follows that

$$[q(t)]_{\mathcal{S}_2} = [T(p(t))]_{\mathcal{S}_2} = [T]_{\mathcal{S}_1}^{\mathcal{S}_2}[p(t)]_{\mathcal{S}_1}.$$



If $p(t) = a_0 + a_1 t$, then

$$\begin{aligned}[q(t)]_{\mathcal{S}_2} &= \begin{bmatrix} 0 & 1 \\ 1 & 0 \\ 2 & 0 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \end{bmatrix} \\ &= \begin{bmatrix} a_1 \\ a_0 \\ 2a_0 \end{bmatrix} \\ &= a_0 \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix} + a_1 \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}.\end{aligned}$$

Thus, $q(t)$ is in the span of $t + 2t^2$ and 1. So $\text{Range}(T) = \text{Span}\{1, t + 2t^2\}$. Since neither 1 nor $t + 2t^2$ is a scalar multiple of the other, the vectors 1 and $t + 2t^2$ are linearly independent. Thus, $\{1, t + 2t^2\}$ is a basis for $\text{Range}(T)$.

(d) Recall that

$$[T]_{\mathcal{B}}^{\mathcal{C}} = [[T(1+t)]_{\mathcal{C}} \ [T(1-t)]_{\mathcal{C}}] = [[1+2t+t^2]_{\mathcal{C}} \ [1-t^2]_{\mathcal{C}}] = \begin{bmatrix} 0 & 0 \\ 1 & 1 \\ 2 & 0 \end{bmatrix}.$$

(e) To find $T(1-2t)$ using the matrix $[T]_{\mathcal{B}}^{\mathcal{C}}$, recall that $[T(p(t))]_{\mathcal{C}} = [T]_{\mathcal{B}}^{\mathcal{C}}[p(t)]_{\mathcal{B}}$. First note that $[1-2t]_{\mathcal{B}} = \begin{bmatrix} -\frac{1}{2} \\ \frac{3}{2} \end{bmatrix}$. So

$$[T(p(t))]_{\mathcal{C}} = [T]_{\mathcal{B}}^{\mathcal{C}}[p(t)]_{\mathcal{B}} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} -\frac{1}{2} \\ \frac{3}{2} \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix}.$$

This makes

$$T(1-2t) = 0(t) + (1)(1-t^2) + (-1)(t+t^2) = 1-t-2t^2.$$

To check, using the definition of T we have

$$T(1-2t) = t(1-2t) + (1-2t) = 1-t-2t^2.$$

Example 36.4. Let $T : V \rightarrow W$ be a linear transformation. Let $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$ be a basis for V and $\mathcal{C} = \{\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_m\}$ a basis for W . Let $S : \mathbb{R}^n \rightarrow \mathbb{R}^m$ be the matrix transformation defined by $S(\mathbf{x}) = [T]_{\mathcal{B}}^{\mathcal{C}}\mathbf{x}$.

- (a) Let \mathbf{w} be a vector in $\text{Range}(T)$. Show that $[\mathbf{w}]_{\mathcal{C}}$ is in $\text{Col } [T]_{\mathcal{B}}^{\mathcal{C}}$.
- (b) If $\mathbf{w} \in \text{Range}(T)$, part (a) shows that $[\mathbf{w}]_{\mathcal{C}}$ is in $\text{Col } [T]_{\mathcal{B}}^{\mathcal{C}}$. So the coordinate transformation $[\]_{\mathcal{C}}$ maps $\text{Range}(T)$ into $\text{Col } [T]_{\mathcal{B}}^{\mathcal{C}}$. Define $R : \text{Range}(T) \rightarrow \text{Col } [T]_{\mathcal{B}}^{\mathcal{C}}$ to be this coordinate transformation. That is, $R(\mathbf{w}) = [\mathbf{w}]_{\mathcal{C}}$. (As a coordinate transformation, R is a linear transformation.) We know that coordinate transformations are one-to-one. Show that R is also an onto transformation.



- (c) What can we conclude about the relationship between the vector spaces $\text{Range}(T)$ and $\text{Col } [T]_{\mathcal{B}}^{\mathcal{C}}$? What must then be true about $\dim(\text{Range}(T))$ and $\dim(\text{Col } [T]_{\mathcal{B}}^{\mathcal{C}})$?

When we connect the results of this example with the result of Exercise 5. in this section, we obtain a linear transformation analog of the Rank-Nullity Theorem. That is, if T is a linear transformation from a vector space V of dimension n with basis \mathcal{B} to a vector space W of dimension m with basis \mathcal{C} , then

$$\dim(\text{Ker}(T)) + \dim(\text{Range}(T)) = \dim(\text{Nul } [T]_{\mathcal{B}}^{\mathcal{C}}) + \dim(\text{Col } [T]_{\mathcal{B}}^{\mathcal{C}}) = n.$$

Example Solution.

- (a) Let \mathbf{w} be a vector in $\text{Range}(T)$. Then there exists a vector \mathbf{v} in V such that $T(\mathbf{v}) = \mathbf{w}$. It follows that

$$[\mathbf{w}]_{\mathcal{C}} = [T(\mathbf{v})]_{\mathcal{C}} = [T]_{\mathcal{B}}^{\mathcal{C}}[\mathbf{v}]_{\mathcal{C}}.$$

Recall that the vectors of the form $A\mathbf{z}$ all linear combinations of the columns of A with weights from \mathbf{z} , so $[T]_{\mathcal{B}}^{\mathcal{C}}[\mathbf{v}]_{\mathcal{C}}$ (or $[\mathbf{w}]_{\mathcal{C}}$) is in $\text{Col } [T]_{\mathcal{B}}^{\mathcal{C}}$.

- (b) Let $R : \text{Range}(T) \rightarrow \text{Col } [T]_{\mathcal{B}}^{\mathcal{C}}$ be the coordinate transformation $R(\mathbf{w}) = [\mathbf{w}]_{\mathcal{C}}$ for each $\mathbf{w} \in \text{Range}(T)$. To show that R is an onto transformation, let \mathbf{y} be in $\text{Col } [T]_{\mathcal{B}}^{\mathcal{C}}$. Then there exists $\mathbf{x} = [x_1 \ x_2 \ \dots \ x_n]^T$ in \mathbb{R}^n such that $[T]_{\mathcal{B}}^{\mathcal{C}}\mathbf{x} = \mathbf{y}$. Let $\mathbf{v} = x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + \dots + x_n\mathbf{v}_n$. Then \mathbf{v} is in V and $[\mathbf{v}]_{\mathcal{B}} = \mathbf{x}$. Also,

$$[T(\mathbf{v})]_{\mathcal{C}} = [T]_{\mathcal{B}}^{\mathcal{C}}[\mathbf{v}]_{\mathcal{B}} = [T]_{\mathcal{B}}^{\mathcal{C}}\mathbf{x} = \mathbf{y}.$$

So if we let $\mathbf{w} = T(\mathbf{v})$ in $\text{Range}(T)$, then $[\mathbf{w}]_{\mathcal{C}} = \mathbf{y}$ and \mathbf{y} and R is onto.

- (c) Since R is a one-to-one and onto linear transformation from $\text{Range}(T)$ to $\text{Col } [T]_{\mathcal{B}}^{\mathcal{C}}$, it follows that $\text{Range}(T)$ and $\text{Col } [T]_{\mathcal{B}}^{\mathcal{C}}$ are isomorphic vector spaces, and therefore we also have $\dim(\text{Range}(T)) = \dim(\text{Col } [T]_{\mathcal{B}}^{\mathcal{C}})$.

Summary

- The standard matrix for a linear transformation T from \mathbb{R}^n to \mathbb{R}^m is the matrix

$$A = [T(\mathbf{e}_1) \ T(\mathbf{e}_2) \ \dots \ T(\mathbf{e}_n)],$$

where $\{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$ is the standard basis for \mathbb{R}^n . Then T is the matrix transformation defined by $T(\mathbf{x}) = A\mathbf{x}$ for all \mathbf{x} in \mathbb{R}^n .

- Let V be an n dimensional vector space and W an m dimensional vector space and suppose $T : V \rightarrow W$ is a linear transformation. Let $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$ be a basis for V and $\mathcal{C} = \{\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_m\}$ a basis for W . The matrix for T with respect to the bases \mathcal{B} and \mathcal{C} is the matrix

$$[T]_{\mathcal{B}}^{\mathcal{C}} = [[T(\mathbf{v}_1)]_{\mathcal{C}} \ [T(\mathbf{v}_2)]_{\mathcal{C}} \ [T(\mathbf{v}_3)]_{\mathcal{C}} \ \dots \ [T(\mathbf{v}_n)]_{\mathcal{C}}].$$

If $V = W$ and $\mathcal{C} = \mathcal{B}$, then we use the notation $[T]_{\mathcal{B}}$ as a shorthand for $[T]_{\mathcal{B}}^{\mathcal{C}}$. The matrix $[T]_{\mathcal{B}}^{\mathcal{C}}$ has the property that

$$[T(\mathbf{v})]_{\mathcal{C}} = [T]_{\mathcal{B}}^{\mathcal{C}}[\mathbf{v}]_{\mathcal{B}}$$

for any \mathbf{v} in V .



- The matrix for a linear transformation allows us to represent any linear transformation between finite dimensional vector spaces as a matrix transformation. When we view a linear transformation T as a matrix transformation, we are able to use the matrix tools we have developed to understand T .

Exercises

(1) Let $T : \mathbb{P}_2 \rightarrow \mathbb{P}_3$ be defined by $T(a_0 + a_1t + a_2t^2) = (a_0 + a_1) + (a_1 + a_2)t + a_0t^2 + a_2t^3$. You may assume that T is a linear transformation. Let $\mathcal{B} = \{p_0(t), p_1(t), p_2(t)\}$, where $p_0(t) = 1$, $p_1(t) = t$, and $p_2(t) = t^2$, be the standard basis for \mathbb{P}_2 and $\mathcal{C} = \{q_0(t), q_1(t), q_2(t), q_3(t)\}$, where $q_0(t) = 1$, $q_1(t) = t$, $q_2(t) = t^2$, and $q_3(t) = t^3$, the standard basis for \mathbb{P}_3 .

- (a) Write the polynomial $r(t) = r_0 + r_1t + r_2t^2$ as a linear combination $c_0p_0(t) + c_1p_1(t) + c_2p_2(t)$ of the basis vectors in \mathcal{B} . Identify the weights c_0 , c_1 , and c_2 . What is $[r(t)]_{\mathcal{B}}$?

- (b) Without doing any calculations, explain why

$$T(r(t)) = c_0T(p_0(t)) + c_1T(p_1(t)) + c_2T(p_2(t)).$$

- (c) Without doing any calculations, explain why

$$[T(r(t))]_{\mathcal{C}} = c_0[T(p_0(t))]_{\mathcal{C}} + c_1[T(p_1(t))]_{\mathcal{C}} + c_2[T(p_2(t))]_{\mathcal{C}}.$$

- (d) Explicitly determine $[T(p_0(t))]_{\mathcal{C}}$, $[T(p_1(t))]_{\mathcal{C}}$, $[T(p_2(t))]_{\mathcal{C}}$.

- (e) Combine the results of parts (c) and (d) to find a matrix A so that

$$[T(r(t))]_{\mathcal{C}} = A[r(t)]_{\mathcal{B}}.$$

- (f) Use the matrix A to find $[T(1+t-t^2)]_{\mathcal{C}}$. Then use this vector to calculate $T(1+t-t^2)$.

- (g) Calculate $T(1+t-t^2)$ directly from the rule for T and compare to the previous result.

(2) Let V and W be finite dimensional vector spaces, and let S and T be linear transformations from V to W .

- (a) The sum $S + T$ is defined as

$$(S + T)(\mathbf{x}) = S(\mathbf{x}) + T(\mathbf{x})$$

for all \mathbf{x} in V . Let \mathcal{B} be a basis for V and \mathcal{C} a basis for W . Is the statement

$$[S + T]_{\mathcal{B}}^{\mathcal{C}} = [S]_{\mathcal{B}}^{\mathcal{C}} + [T]_{\mathcal{B}}^{\mathcal{C}}$$

true or false? If true, prove the statement. If false, provide a counterexample.

- (b) The scalar multiple cT , for a scalar c , is defined as

$$(cT)(\mathbf{x}) = cT(\mathbf{x})$$

for all \mathbf{x} in V . Let \mathcal{B} be a basis for V and \mathcal{C} a basis for W . Is the statement

$$[cT]_{\mathcal{B}}^{\mathcal{C}} = c[T]_{\mathcal{B}}^{\mathcal{C}}$$

true or false? If true, prove the statement. If false, provide a counterexample.



- (3) If V , W , and U are finite dimensional vector spaces and $S : V \rightarrow W$ and $T : W \rightarrow U$ are linear transformations, the composite $T \circ S$ is defined as

$$(T \circ S)(\mathbf{x}) = T(S(\mathbf{x}))$$

for all \mathbf{x} in V .

- (a) Prove that $T \circ S : V \rightarrow U$ is a linear transformation.
 (b) Let \mathcal{B} be a basis for V , \mathcal{C} a basis for W , and \mathcal{D} a basis for U . Is the statement

$$[T \circ S]_{\mathcal{B}}^{\mathcal{D}} = [T]_{\mathcal{C}}^{\mathcal{D}} [S]_{\mathcal{B}}^{\mathcal{C}}$$

true or false? If true, prove the statement. If false, provide a counterexample.

- (4) Let V be a finite dimensional vector space with basis \mathcal{B} , and let T be a one-to-one linear transformation from V to V .

- (a) Use the matrix $[T]_{\mathcal{B}}$ to explain why T is also onto. (Recall that we use the shorthand notation $[T]_{\mathcal{B}}$ for $[T]_{\mathcal{B}}^{\mathcal{B}}$.)
 (b) Since T is one-to-one and onto, as a function T has an inverse defined by

$$T^{-1}(\mathbf{y}) = \mathbf{x}$$

whenever $T(\mathbf{x}) = \mathbf{y}$. Show that T^{-1} is a linear transformation from V to V .

- (c) Is the statement

$$[T^{-1}]_{\mathcal{B}} = ([T]_{\mathcal{B}})^{-1}$$

true or false? If true, prove the statement. If false, provide a counterexample.

- (5) Let V and W be vector spaces with $\dim(V) = n$ and $\dim(W) = m$, and let $T : V \rightarrow W$. Let \mathcal{B} be a basis for V and \mathcal{C} a basis for W . There is a connection between $\text{Ker}(T)$ and $\text{Nul } [T]_{\mathcal{B}}^{\mathcal{C}}$. Find the connection and verify it.

- (6) Let V and W be vector spaces and define $L(V, W)$ to be the set of all linear transformations from V to W as in Exercise 14 of Section 35. The set $L(V, W)$ is a vector space with the operations as follows for S and T are in $L(V, W)$ and c a scalar:

$$(S + T)(\mathbf{v}) = S(\mathbf{v}) + T(\mathbf{v}) \quad \text{and} \quad (cT)(\mathbf{v}) = cT(\mathbf{v})$$

for all \mathbf{v} in V . If $\dim(V) = n$ and $\dim(W) = m$, find the dimension of $L(V, W)$. (Hint: Let \mathcal{B} be a basis for V and \mathcal{C} a basis for W . What can be said about the mapping that sends T in $L(V, W)$ to $[T]_{\mathcal{B}}^{\mathcal{C}}$? Then use Exercise 7. in Section 35.)

- (7) Let $T : V \rightarrow W$ be a linear transformation from an n -dimensional vector space V to an m -dimensional vector space W . Let \mathcal{B} be a basis for V and \mathcal{C} a basis for W . Let $A = [T]_{\mathcal{B}}^{\mathcal{C}}$. Let T' be the matrix transformation from \mathbb{R}^n to \mathbb{R}^m defined by $T'(\mathbf{x}) = A\mathbf{x}$.

- (a) Show that if \mathbf{w} is in $\text{Range}(T)$, then $[\mathbf{w}]_{\mathcal{C}}$ is in the range of T' .
 (b) Let $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$ and $\mathcal{C} = \{\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_m\}$. Show that if the vector $\mathbf{y} = [y_1 \ y_2 \ \dots \ y_m]$ is in the range of T' , then the vector $\mathbf{w} = y_1\mathbf{w}_1 + y_2\mathbf{w}_2 + \dots + y_m\mathbf{w}_m$ is in the range of T .



- (c) Explain why the results of (a) and (b) show that T is onto if and only if every row of $A = [T]_{\mathcal{B}}^{\mathcal{C}}$ contains a pivot.
- (8) Our entire world is always in a state of motion. Much of this motion is vibration. When the vibrations of one object come into contact with the vibrations of another, the vibrations can be amplified. This is called *resonance*. Resonance appears in our lives every day, e.g., resonance is how food cooks in a microwave oven. A great example of resonance is the collapse of the Tacoma Narrows Bridge (just Google Tacoma Narrows Bridge to find some fascinating videos of this event). Resonance can be seen in mathematical models of vibrations. One such model is the second order differential equation

$$y'' + y = \cos(t),$$

where y is some unknown function of t and y'' is its second derivative. You can think of this system as modeling the oscillatory motion of a spring with mass attached to it. The cosine function in this example exerts an external force on the mass-spring system. (You don't need to know how this model is derived for this project.) Our goal is to find the solutions to this differential equation within the subspace $Y = \text{Span } \mathcal{B}$, where $\mathcal{B} = \{\cos(t), \sin(t), t \cos(t), t \sin(t)\}$ of \mathbb{D}^2 , the subspace of \mathcal{F} of functions with second derivatives (you may assume that \mathbb{D}^2 is a subspace of \mathcal{F}).

Let $T : \mathbb{D}^2 \rightarrow \mathbb{D}^2$ be defined by $T(f) = f'' + f$.

- (a) Show that T is a linear transformation.
- (b) Show that \mathcal{B} is a basis for Y . (Hint: You might consider using the Wronskian.)
- (c) Find the matrix of T with respect to the basis \mathcal{B} . That is, find $[T]_{\mathcal{B}}$. (Recall that we use the shorthand notation $[T]_{\mathcal{B}}$ for $[T]_{\mathcal{B}}^{\mathcal{B}}$.)
- (d) Use the matrix to find *all* solutions to the equation $T(f) = \cos(t)$ in Y . (Hint: If $T(f) = \cos(t)$, what does that say about the relationship between $[T]_{\mathcal{B}}^{\mathcal{B}}$, $[\cos(t)]_{\mathcal{B}}$, and $[f]_{\mathcal{B}}$?)
- (e) Sketch a few graphs of solutions to $T(f) = \cos(t)$ and explain what they look like and how they are related to resonance.
- (9) Label each of the following statements as True or False. Provide justification for your response.
- (a) **True/False** If T is a linear transformation from a vector space V to a vector space W , and $[T]_{\mathcal{B}}^{\mathcal{C}}$ is a 3×2 matrix for some bases \mathcal{B} of V and \mathcal{C} of W , then T cannot be one-to-one.
- (b) **True/False** If T is a linear transformation from a vector space V to a vector space W , and $[T]_{\mathcal{B}}^{\mathcal{C}}$ is a 2×3 matrix for some bases \mathcal{B} of V and \mathcal{C} of W , then T cannot be onto.
- (c) **True/False** If T is a linear transformation from a vector space V to a vector space W , and $[T]_{\mathcal{B}}^{\mathcal{C}}$ is a 2×3 matrix for some bases \mathcal{B} of V and \mathcal{C} of W , then T cannot be one-to-one.



- (d) **True/False** If T is a linear transformation from a vector space V to a vector space W , and $[T]_{\mathcal{B}}^{\mathcal{C}}$ is a 3×2 matrix for some bases \mathcal{B} of V and \mathcal{C} of W , then T cannot be onto.
- (e) **True/False** Let S be a linear transformation from a vector space Y to a vector space V and let T be a linear transformation from a vector space V to a vector space W . Suppose that $[S]_{\mathcal{S}}^{\mathcal{B}}$ is a 3×4 matrix and $[T]_{\mathcal{B}}^{\mathcal{C}}$ is a 4×2 matrix for some bases \mathcal{S} of Y , \mathcal{B} of V and \mathcal{C} of W . Then $[T \circ S]_{\mathcal{S}}^{\mathcal{C}}$ is a 3×2 matrix.
- (f) **True/False** Let V be a finite dimensional vector space of dimension n with basis \mathcal{B} and W a finite dimensional vector space of dimension m with basis \mathcal{C} . Let T be a linear transformation from V to W . If the columns of the matrix $[T]_{\mathcal{B}}^{\mathcal{C}}$ are linearly independent, then the transformation T is onto.
- (g) **True/False** Let V be a finite dimensional vector space of dimension n with basis \mathcal{B} and W a finite dimensional vector space of dimension m with basis \mathcal{C} . Let T be a linear transformation from V to W . If the columns of the matrix $[T]_{\mathcal{B}}^{\mathcal{C}}$ are linearly independent, then the transformation T is one-to-one.
- (h) **True/False** Let V be a finite dimensional vector space of dimension n with basis \mathcal{B} and W a finite dimensional vector space of dimension m with basis \mathcal{C} . Let T be a linear transformation from V to W . If the matrix $[T]_{\mathcal{B}}^{\mathcal{C}}$ has n pivots, then the transformation T is onto.
- (i) **True/False** Let V be a finite dimensional vector space of dimension n with basis \mathcal{B} and W a finite dimensional vector space of dimension m with basis \mathcal{C} . Let T be a linear transformation from V to W . If the matrix $[T]_{\mathcal{B}}^{\mathcal{C}}$ has n pivots, then the transformation T is one-to-one.

Project: Shamir's Secret Sharing and Lagrange Polynomials

Shamir's Secret Sharing is a secret sharing algorithm developed by the Israeli cryptographer Adi Shamir, who also contributed to the invention of RSA algorithm. The idea is to develop shares related to a secret code that can be distributed in such a way that the secret code can only be discovered if a fixed number of shareholders combine their information.

The general algorithm for Shamir's Secret Sharing works by creating a polynomial with the secret code as one coefficient and random remaining coefficients. More specifically, to create a (k, n) threshold scheme, let a_0 be the secret code. Then choose at random $k - 1$ positive integers a_1, a_2, \dots, a_{k-1} and create the polynomial

$$p(t) = a_0 + a_1 t + a_2 t^2 + \cdots + a_{k-1} t^{k-1},$$

where a_0 is the secret. Evaluate $p(t)$ at n different inputs t_1, t_2, \dots, t_n to create n points $P_1 = (t_1, p(t_1)), P_2 = (t_2, p(t_2)), \dots, P_n = (t_n, p(t_n))$. Each participant is given one of these points. Since any collection of k distinct points on the graph of $p(t)$ uniquely determines $p(t)$, any combination of k of the participants could reconstruct $p(t)$. The secret is then $p(0)$. (Note that, except under very restrictive circumstances, there is no polynomial of degree less than k that passes through the given k points, so it would be impossible for fewer than k participants to reconstruct the message.)



As an example, let our secret code be $a_0 = 1234$. Let $n = 6$ and $k = 3$. Choose two positive integers at random, say $a_1 = 100$ and $a_2 = 38$. Then

$$p(t) = 1234 + 100t + 38t^2. \quad (36.2)$$

Now we construct 6 points by selecting 6 positive integers, say $t_1 = 1, t_2 = 2, t_3 = 3, t_4 = 4, t_5 = 5$, and $t_6 = 6$. Evaluating $p(t)$ at $t = 1$ through $t = 6$ gives us the points $P_1 = (1, 1372), P_2 = (2, 1586), P_3 = (3, 1876), P_4 = (4, 2242), P_5 = (5, 2684)$, and $P_6 = (6, 3202)$. Our goal is to find the polynomial that passes through any three of these points. The next activity will show us how to find this polynomial.

Project Activity 36.1. It will be instructive to work in the most general setting. We restrict ourselves to the degree 2 case for now. Given three scalars t_1, t_2 , and t_3 , define $L : \mathbb{P}_2 \rightarrow \mathbb{R}^3$ by

$$L(q(t)) = \begin{bmatrix} q(t_1) \\ q(t_2) \\ q(t_3) \end{bmatrix}.$$

- (a) Show that L is a linear transformation.
- (b) Find the matrix for L with respect to the standard bases $\mathcal{B} = \{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ for \mathbb{R}^3 and $\mathcal{S} = \{1, t, t^2\}$ for \mathbb{P}_2 .
- (c) Recall that the matrix $[L]_{\mathcal{S}}^{\mathcal{B}}$ has the property that $[L(q(t))]_{\mathcal{B}} = [L]_{\mathcal{S}}^{\mathcal{B}}[q(t)]_{\mathcal{S}}$ for any $q(t)$ in \mathbb{P}_2 . So L is one-to-one if and only if the matrix transformation defined by $[L]_{\mathcal{S}}^{\mathcal{B}}$ is one-to-one. Use this idea to show that L is one-to-one if and only if t_1, t_2 , and t_3 are all different. Use appropriate technology where needed.
- (d) Given three points $(t_1, y_1), (t_2, y_2)$, and (t_3, y_3) with distinct t_1, t_2 , and t_3 , our objective is to find a polynomial $p(t)$ such that $p(t_i) = y_i$ for i from 1 to 3. We proceed by finding quadratic polynomials $\ell_1(t), \ell_2(t)$, and $\ell_3(t)$ so that $L(\ell_1(t)) = \mathbf{e}_1, L(\ell_2(t)) = \mathbf{e}_2$, and $L(\ell_3(t)) = \mathbf{e}_3$. Explain why, if $\ell_1(t), \ell_2(t)$, and $\ell_3(t)$ are quadratic polynomials so that $L(\ell_1(1)) = \mathbf{e}_1, L(\ell_2(t)) = \mathbf{e}_2$, and $L(\ell_3(t)) = \mathbf{e}_3$, and t_1, t_2 , and t_3 are all different, then the polynomial $p(t)$ will satisfy

$$p(t) = y_1\ell_1(t) + y_2\ell_2(t) + y_3\ell_3(t). \quad (36.3)$$

The polynomials $\ell_1(t), \ell_2(t)$, and $\ell_3(t)$ in Project Activity 36.1 are examples of *Lagrange polynomials*. Project Activity 36.1 shows that fitting polynomials to given points can be accomplished easily using linear combinations of Lagrange polynomials. Now we want to better understand the general formulas for the Lagrange polynomials $\ell_i(t)$.

Project Activity 36.2. In this activity we see how to find the quadratic polynomials $\ell_1(t), \ell_2(t)$, and $\ell_3(t)$ so that $L(\ell_1(t)) = \mathbf{e}_1, L(\ell_2(t)) = \mathbf{e}_2$, and $L(\ell_3(t)) = \mathbf{e}_3$. Let $\mathcal{B} = \{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ and $\mathcal{S} = \{1, t, t^2\}$ be the standard bases for \mathbb{R}^3 and \mathbb{P}_2 , respectively.

- (a) Explain why the problem of solving the equations $L(\ell_1(t)) = \mathbf{e}_1, L(\ell_2(t)) = \mathbf{e}_2$, and $L(\ell_3(t)) = \mathbf{e}_3$ is equivalent to solving the equations

$$[L]_{\mathcal{S}}^{\mathcal{B}}[\ell_1(t)]_{\mathcal{S}} = \mathbf{e}_1, [L]_{\mathcal{S}}^{\mathcal{B}}[\ell_2(t)]_{\mathcal{S}} = \mathbf{e}_2, \text{ and } [L]_{\mathcal{S}}^{\mathcal{B}}[\ell_3(t)]_{\mathcal{S}} = \mathbf{e}_3.$$



- (b) Explain why we can solve the equations

$$[L]_{\mathcal{S}}^{\mathcal{B}}[\ell_1(t)]_{\mathcal{S}} = \mathbf{e}_1, [L]_{\mathcal{S}}^{\mathcal{B}}[\ell_2(t)]_{\mathcal{S}} = \mathbf{e}_2, \text{ and } [L]_{\mathcal{S}}^{\mathcal{B}}[\ell_3(t)]_{\mathcal{S}} = \mathbf{e}_3$$

all at once by solving the matrix equation

$$[L]_{\mathcal{S}}^{\mathcal{B}} [[\ell_1(t)]_{\mathcal{S}} \ [\ell_2(t)]_{\mathcal{S}} \ [\ell_3(t)]_{\mathcal{S}}] = I_3.$$

What does this tell us about the relationship between the matrix $[[\ell_1(t)]_{\mathcal{S}} \ [\ell_2(t)]_{\mathcal{S}} \ [\ell_3(t)]_{\mathcal{S}}]$ and $[L]_{\mathcal{S}}^{\mathcal{B}}$?

- (c) Technology shows that

$$([L]_{\mathcal{S}}^{\mathcal{B}})^{-1} = \begin{bmatrix} \frac{t_2 t_3}{(t_1 - t_2)(t_1 - t_3)} & \frac{t_1 t_3}{(t_2 - t_1)(t_2 - t_3)} & \frac{t_1 t_2}{(t_3 - t_1)(t_3 - t_2)} \\ -\frac{t_2 + t_3}{(t_1 - t_2)(t_1 - t_3)} & -\frac{t_1 + t_3}{(t_2 - t_1)(t_2 - t_3)} & -\frac{t_1 + t_2}{(t_3 - t_1)(t_3 - t_2)} \\ \frac{1}{(t_1 - t_2)(t_1 - t_3)} & \frac{1}{(t_2 - t_1)(t_2 - t_3)} & \frac{1}{(t_3 - t_1)(t_3 - t_2)} \end{bmatrix}.$$

Use $([L]_{\mathcal{S}}^{\mathcal{B}})^{-1}$ to determine the coefficients of $\ell_1(t)$. Then apply some algebra to show that

$$\ell_1(t) = \frac{(t - t_2)(t - t_3)}{(t_1 - t_2)(t_1 - t_3)}.$$

- (d) Find similar expressions for $\ell_2(t)$ and $\ell_3(t)$. Explain why these three polynomials satisfy the required conditions that $L(\ell_1(t)) = \mathbf{e}_1$, $L(\ell_2(t)) = \mathbf{e}_2$, and $L(\ell_3(t)) = \mathbf{e}_3$.

Now we return to our secret code with $a_0 = 1234$.

Project Activity 36.3. Pick any three of the points $P_1 = (1, 1372)$, $P_2 = (2, 1586)$, $P_3 = (3, 1876)$, $P_4 = (4, 2242)$, $P_5 = (5, 2684)$, and $P_6 = (6, 3202)$. Use the Lagrange polynomials $\ell_1(t)$, $\ell_2(t)$, and $\ell_3(t)$ and (36.3) to find the polynomial whose graph contains the three chosen points. How does your polynomial compare to the polynomial $p(t)$ in (36.2)?

The Shamir Secret Sharing algorithm depends on being able to find a unique polynomial $p(t)$ that passes through the created points. We can understand this result with Lagrange polynomials.

Project Activity 36.4. The process described in Project Activity 36.2 for finding the Lagrange polynomials can be applied to any number of points. Let (t_1, y_1) , (t_2, y_2) , (t_3, y_3) , \dots , (t_n, y_n) , and (t_{n+1}, y_{n+1}) be $n + 1$ points with distinct t_i . Generalizing the results of Project Activity 36.2, define $\ell_i(t)$ for i from 1 to $n + 1$ as follows:

$$\begin{aligned} \ell_i(t) &= \frac{(t - t_1)(t - t_2) \cdots (t - t_{i-1})(t - t_{i+1}) \cdots (t - t_n)(t - t_{n+1})}{(t_i - t_1)(t_i - t_2) \cdots (t_i - t_{i-1})(t_i - t_{i+1}) \cdots (t_i - t_n)(t_i - t_{n+1})} \\ &= \prod_{j \neq i} \frac{t - t_j}{t_i - t_j}. \end{aligned}$$

- (a) Explain why $\ell_i(t)$ is in \mathbb{P}_n for each i .



(b) Explain why the polynomial $\ell_i(t)$ satisfies $\ell_i(t_i) = 1$ and $\ell_i(t_j) = 0$ if $i \neq j$.

(c) Let $p(t) = \sum_{i=1}^{n+1} y_i \ell_i(t)$. That is, $p(t)$ is the polynomial

$$\sum_{i=1}^{n+1} y_i \frac{(t - t_1)(t - t_2) \dots (t - t_{i-1})(t - t_{i+1}) \dots (t - t_n)(t - t_{n+1})}{(t_i - t_1)(t_i - t_2) \dots (t_i - t_{i-1})(t_i - t_{i+1}) \dots (t_i - t_n)(t_i - t_{n+1})}.$$

Explain why $p(t)$ is a polynomial in \mathbb{P}_n whose graph contains the points (t_i, y_i) for i from 1 to $n + 1$.

(d) The previous part demonstrates that there is always a polynomial in \mathbb{P}_n whose graph contains the points $(t_1, y_1), (t_2, y_2), (t_3, y_3), \dots, (t_n, y_n)$, and (t_{n+1}, y_{n+1}) with distinct t_i . The last piece we need is to know that this polynomial is unique. Use the fact that a polynomial of degree n can have at most n roots to show that if $f(t)$ and $g(t)$ are two polynomials in \mathbb{P}_n whose graphs contain the points $(t_1, y_1), (t_2, y_2), \dots, (t_n, y_n), (t_{n+1}, y_{n+1})$ with distinct values of t_i , then $f(t) = g(t)$. This completes Shamir's Secret Sharing algorithm.

Section 37

Eigenvalues of Linear Transformations

Focus Questions

By the end of this section, you should be able to give precise and thorough answers to the questions listed below. You may want to keep these questions in mind to focus your thoughts as you complete the section.

- How, and under what conditions, do we define the eigenvalues and eigenvectors of a linear transformation?
- How do we find eigenvalues and eigenvectors of a linear transformation? What important result provides that our method for doing so works?
- What does it mean to diagonalize a linear transformation?
- Under what conditions is a linear transformation diagonalizable?

Application: Linear Differential Equations

A body in motion obeys Newton's second law that force equals mass times acceleration, or $F = ma$. Here F is the force acting on the object, m the mass of the object, and a the acceleration of the object. For example, if a mass is hanging from a spring, gravity acts to pull the mass downward and the spring acts to pull the mass up. Hooke's law says that the force of the spring acting on mass is proportional to the displacement y of the spring from equilibrium. There is also a damping force that weakens the action of the spring that might be due to air resistance or the medium in which the system is enclosed. If this mass is not too large, then the resistance can be taken to be proportional to the velocity v of the mass. This produces forces $F = ma$ acting to pull the mass down and forces $-ky$ and $-bv$ acting in the opposite direction, where k and b are constant. A similar type of analysis would apply to an object falling through the air, where the resistant to falling could be proportional to the velocity of the object. Since $v = y'$ and $a = y''$, equating the forces produces the equation'

$$my'' = -ky - bv. \quad (37.1)$$

Any equation that involves one or more derivatives of a function is called a *differential equation*. Differential equations are used to model many different types of behavior in disciplines including engineering, physics, economics, and biology. Later in this section we will see how differential equations can be represented by linear transformations, and see how we can exploit this idea to solve certain types of differential equations (including (37.1)).

Introduction

Recall that a scalar λ is an eigenvalue of an $n \times n$ matrix A if $A\mathbf{x} = \lambda\mathbf{x}$ for some nonzero vector \mathbf{x} in \mathbb{R}^n . Now that we have seen how to represent a linear transformation T from a finite dimensional vector space V to itself with a matrix transformation, we can exploit this idea to define and find eigenvalues and eigenvectors for T just as we did for matrices.

Definition 37.1. Let V be an n dimensional vector space and $T : V \rightarrow V$ a linear transformation. A scalar λ is an **eigenvalue** for T if there is a nonzero vector \mathbf{v} in V such that $T(\mathbf{v}) = \lambda\mathbf{v}$. The vector \mathbf{v} is an **eigenvector** for T corresponding to the eigenvalue λ .

We can exploit the fact that we can represent linear transformations as matrix transformations to find eigenvalues and eigenvectors of a linear transformation.

Preview Activity 37.1. Let $T : \mathbb{P}_1 \rightarrow \mathbb{P}_1$ be defined by $T(a_0 + a_1t) = (a_0 + 2a_1) + (3a_0 + 2a_1)t$. Assume T is a linear transformation.

- (1) Let $\mathcal{S} = \{1, t\}$ be the standard basis for \mathbb{P}_1 . Find the matrix $[T]_{\mathcal{S}}$. (Recall that we use the shorthand notation $[T]_{\mathcal{S}}$ for $[T]_{\mathcal{S}}^{\mathcal{S}}$.)
- (2) Check that $\lambda_1 = 4$ and $\lambda_2 = -1$ are the eigenvalues of $[T]_{\mathcal{S}}$. Find an eigenvector \mathbf{v}_1 for $[T]_{\mathcal{S}}$ corresponding to the eigenvalue 4 and an eigenvector \mathbf{v}_2 of $[T]_{\mathcal{S}}$ corresponding to the eigenvalue -1 .
- (3) Find the vector in \mathbb{P}_1 corresponding to the eigenvector of $[T]_{\mathcal{S}}$ for the eigenvalue $\lambda = 4$. Check that this is an eigenvector of T .
- (4) Explain why in general, if V is a vector space with basis \mathcal{B} and $S : V \rightarrow V$ is a linear transformation, and if \mathbf{v} in V satisfies $[\mathbf{v}]_{\mathcal{B}} = \mathbf{w}$, where \mathbf{w} is an eigenvector of $[S]_{\mathcal{B}}$ with eigenvalue λ , then \mathbf{v} is an eigenvector of T with eigenvalue λ .

Finding Eigenvalues and Eigenvectors of Linear Transformations

Preview Activity 37.1 presents a method for finding eigenvalues and eigenvectors of linear transformations. That is, if $T : V \rightarrow V$ is a linear transformation and \mathcal{B} is a basis for V , and if \mathbf{x} is an eigenvector of $[T]_{\mathcal{B}}$ with eigenvalue λ , then the vector \mathbf{v} in V satisfying $[\mathbf{v}]_{\mathcal{B}} = \mathbf{x}$ has the property that

$$[T(\mathbf{v})]_{\mathcal{B}} = [T]_{\mathcal{B}}[\mathbf{v}]_{\mathcal{B}} = [T]_{\mathcal{B}}\mathbf{x} = \lambda\mathbf{x} = [\lambda\mathbf{v}]_{\mathcal{B}}.$$

Since the coordinate transformation is one-to-one, it follows that $T(\mathbf{v}) = \lambda\mathbf{v}$ and \mathbf{v} is an eigenvector of T with eigenvalue λ . So every eigenvector of $[T]_{\mathcal{B}}$ corresponds to an eigenvector of T . The fact



that every eigenvector and eigenvalue of T can be obtained from the eigenvectors and eigenvalues of $[T]_{\mathcal{B}}$ is left as Exercise 2. Now we address the question of how eigenvalues of matrices of T with respect to different bases are related.

Activity 37.1. Let $T : \mathbb{P}_1 \rightarrow \mathbb{P}_1$ still be defined by $T(a_0 + a_1t) = (a_0 + 2a_1) + (2a_1 + 3a_0)t$.

- Let $\mathcal{B} = \{1 + t, 1 - t\}$ be a basis for \mathbb{P}_1 . Find the matrix $[T]_{\mathcal{B}}$.
- Check that $\lambda_1 = 4$ and $\lambda_2 = -1$ are the eigenvalues of $[T]_{\mathcal{B}}$. Find an eigenvector \mathbf{w}_1 for $[T]_{\mathcal{B}}$ corresponding to the eigenvalue 4.
- Use the matrix $[T]_{\mathcal{B}}$ to find an eigenvector for T corresponding to the eigenvalue $\lambda_1 = 4$.

You should notice that the matrices $[T]_{\mathcal{S}}$ and $[T]_{\mathcal{B}}$ in Preview Activity 37.1 and Activity 37.1 are similar matrices. In fact, if $P = \begin{matrix} P \\ \mathcal{B} \leftarrow \mathcal{S} \end{matrix}$ is the change of basis matrix from \mathcal{S} to \mathcal{B} , then $P^{-1}[T]_{\mathcal{B}}P = [T]_{\mathcal{S}}$. So it must be the case that $[T]_{\mathcal{S}}$ and $[T]_{\mathcal{B}}$ have the same eigenvalues. What Preview Activity 37.1, Exercise 2, and Activity 37.1 demonstrate is that the eigenvalues of a linear transformation $T : V \rightarrow V$ for any n dimensional vector space V are the eigenvalues of the matrix for T relative to any basis for V . We can then find corresponding eigenvectors for T using the methods demonstrated in Preview Activity 37.1 and Activity 37.1. That is, if \mathbf{v} is an eigenvector of T with eigenvalue λ , and \mathcal{B} is any basis for V , then

$$[T]_{\mathcal{B}}[\mathbf{v}]_{\mathcal{B}} = [T(\mathbf{v})]_{\mathcal{B}} = [\lambda\mathbf{v}]_{\mathcal{B}} = \lambda[\mathbf{v}]_{\mathcal{B}}.$$

Thus, once we find an eigenvector $[\mathbf{v}]_{\mathcal{B}}$ for $[T]_{\mathcal{B}}$, the vector \mathbf{v} is an eigenvector for T . This is summarized in the following theorem.

Theorem 37.2. *Let V be a finite dimensional vector space with basis \mathcal{B} and let $T : V \rightarrow V$ be a linear transformation. A vector \mathbf{v} in V is an eigenvector of T with eigenvalue λ if and only if the vector $[\mathbf{v}]_{\mathcal{B}}$ is an eigenvector of $[T]_{\mathcal{B}}$ with eigenvalue λ .*

Diagonalization

If $T : V \rightarrow V$ is a linear transformation, one important question is that of finding a basis for V in which the matrix of T has as simple a form as possible. In other words, can we find a basis \mathcal{B} for V for which the matrix $[T]_{\mathcal{B}}$ is a diagonal matrix? Since any matrices for T with respect to different bases are similar, this will happen if we can diagonalize any matrix for T .

Definition 37.3. Let V be a vector space of dimension n . A linear transformation T from V to V is **diagonalizable** if there is a basis \mathcal{B} for V for which $[T]_{\mathcal{B}}$ is a diagonalizable matrix.

Now the question is, if T is diagonalizable, how do we find a basis \mathcal{B} to that $[T]_{\mathcal{B}}$ is a diagonal matrix? Recall that to diagonalize a matrix A means to find a matrix P so that $P^{-1}AP$ is a diagonal matrix.

Activity 37.2. Let $V = \mathbb{P}_1$, $T : V \rightarrow V$ defined by $T(a_0 + a_1t) = (a_0 + 2a_1) + (2a_1 + a_0)t$. Let $\mathcal{S} = \{1, t\}$ be the standard basis for V .

- Find the matrix $[T]_{\mathcal{S}}$ for T relative to the basis \mathcal{S} .



- (b) Use the fact that the eigenvalues of $[T]_S$ are -1 and 3 with corresponding eigenvectors $\begin{bmatrix} -1 \\ 1 \end{bmatrix}$ and $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$, respectively, to find a matrix P so that $P^{-1}[T]_SP$ is a diagonal matrix D .
- (c) To find a basis \mathcal{B} for which $[T]_{\mathcal{B}}$ is diagonal, note that if $P = \begin{smallmatrix} P \\ \mathcal{B} \leftarrow S \end{smallmatrix}$ of the previous part is the change of basis matrix from S to \mathcal{B} , then the matrices of T with respect to S and \mathcal{B} are related by $P^{-1}[T]_SP = [T]_{\mathcal{B}}$, which makes $[T]_{\mathcal{B}}$ a diagonal matrix. Use the fact that $P[\mathbf{x}]_{\mathcal{B}} = [\mathbf{x}]_S$ to find the vectors in the basis \mathcal{B} .
- (d) Now show directly that $[T]_{\mathcal{B}} = D$, where $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2\}$, and verify that we have found a basis for V for which the matrix for T is diagonal.

The general idea for diagonalizing a linear transformation is contained in Activity 37.2. Let V be an n dimensional vector space and assume the linear transformation $T : V \rightarrow V$ is diagonalizable. So there exists a basis $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$ for which $[T]_{\mathcal{B}}$ is a diagonal matrix. To find this basis, note that for any other basis \mathcal{C} we know that $[T]_{\mathcal{C}}$ and $[T]_{\mathcal{B}}$ are similar. That means that there is an invertible matrix P so that $P^{-1}[T]_{\mathcal{C}}P = [T]_{\mathcal{B}}$, where $P = \begin{smallmatrix} P \\ \mathcal{C} \leftarrow \mathcal{B} \end{smallmatrix}$ is the change of basis matrix from \mathcal{B} to \mathcal{C} . So to find \mathcal{B} , we choose \mathcal{C} so that P is the matrix that diagonalizes $[T]_{\mathcal{C}}$. Using the definition of the change of basis matrix, we then know that each basis vector \mathbf{v}_i in \mathcal{B} satisfies $[\mathbf{v}_i]_{\mathcal{C}} = P[\mathbf{v}_i]_{\mathcal{B}} = P\mathbf{e}_i$. From this we can find \mathbf{v}_i . Note that a standard basis S is often a convenient choice to make for the basis \mathcal{C} . This is the process we used in Activity 37.2.

Recall that an $n \times n$ matrix A is diagonalizable if and only if A has n linearly independent eigenvectors. We apply this idea to a linear transformation as well, summarized by the following theorem.

Theorem 37.4. *Let V be a vector space of dimension n and let T be a linear transformation from V to V . Then the following are equivalent.*

- (1) T is diagonalizable.
- (2) There exists a basis \mathcal{B} for V and an invertible matrix P so that $P^{-1}[T]_{\mathcal{B}}P$ is a diagonal matrix.
- (3) There exists a basis \mathcal{C} of V for which the matrix $[T]_{\mathcal{C}}$ has n linearly independent eigenvectors.
- (4) There exists a basis of V consisting of eigenvectors of T .

Examples

What follows are worked examples that use the concepts from this section.

Example 37.5. Let $T : \mathcal{M}_{2 \times 2} \rightarrow \mathcal{M}_{2 \times 2}$ be defined by $T(A) = A^T$. Let $M_1 = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}$, $M_2 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$, $M_3 = \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}$, and $M_4 = \begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix}$, and let \mathcal{B} be the ordered basis $\{M_1, M_2, M_3, M_4\}$.



- (a) Show that T is a linear transformation.
- (b) Find $[T]_{\mathcal{B}}$.
- (c) Is the matrix $[T]_{\mathcal{B}}$ diagonalizable? If so, find a matrix P such that $P^{-1}[T]_{\mathcal{B}}P$ is a diagonal matrix. Use technology as appropriate.
- (d) Use part (c) to find a basis \mathcal{C} for $\mathcal{M}_{2 \times 2}$ for which $[T]_{\mathcal{C}}$ is a diagonal matrix.

Example Solution.

- (a) Let A and B be 2×2 matrices and let c be a scalar. The properties of the transpose show that

$$T(A + B) = (A + B)^T = A^T + B^T = T(A) + T(B)$$

and

$$T(cA) = (cA)^T = cA^T = cT(A).$$

Thus, T is a linear transformation.

- (b) Notice that $T(M_1) = M_3$, $T(M_2) = M_2$, $T(M_3) = M_1$, and $T(M_4) = M_1 - M_3 + M_4$. So

$$[T]_{\mathcal{B}} = [[T(M_1)]_{\mathcal{B}} \ [T(M_2)]_{\mathcal{B}} \ [T(M_3)]_{\mathcal{B}} \ [T(M_4)]_{\mathcal{B}}] = \begin{bmatrix} 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & -1 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

- (c) Technology shows that the characteristic polynomial of $[T]_{\mathcal{B}}$ is $p(\lambda) = (\lambda + 1)(\lambda - 1)^3$.

It follows that the eigenvalues of $[T]_{\mathcal{B}}$ are -1 and 1 (with multiplicity 3). Technology also shows that $\{[-1 \ 0 \ 1 \ 0]^T\}$ is a basis for the eigenspace corresponding to the eigenvalue -1 and the vectors $[1 \ 0 \ 0 \ 1]^T$, $[1 \ 0 \ 1 \ 0]^T$, $[0 \ 1 \ 0 \ 0]^T$ form a basis for the eigenspace

corresponding to the eigenvalue 1 . If we let $P = \begin{bmatrix} 1 & 1 & 0 & -1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{bmatrix}$, then $P^{-1}AP =$

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}.$$

- (d) Suppose $\mathcal{C} = \{Q_1, Q_2, Q_3, Q_4\}$ is a basis for $\mathcal{M}_{2 \times 2}$ satisfying $P^{-1}[T]_{\mathcal{B}}P = [T]_{\mathcal{C}}$. This makes $[T]_{\mathcal{C}}$ a diagonal matrix with the eigenvalues of $[T]_{\mathcal{B}}$ along the diagonal. In this case we have

$$[T]_{\mathcal{C}}[A]_{\mathcal{C}} = P^{-1}[T]_{\mathcal{B}}P[A]_{\mathcal{C}}$$

for any matrix $A \in \mathcal{M}_{2 \times 2}$. So P is the change of basis matrix from \mathcal{C} to \mathcal{B} . That is, $P[A]_{\mathcal{C}} = [A]_{\mathcal{B}}$. It follows that $P[Q_i]_{\mathcal{C}} = [Q_i]_{\mathcal{B}}$ for i from 1 to 4. Since $[Q_i]_{\mathcal{C}} = \mathbf{e}_i$, we can see that $[Q_i]_{\mathcal{B}}$ is the i th column of P . So the columns of P provide the weights for basis vectors in \mathcal{C} in terms of the basis \mathcal{B} . Letting $Q_1 = M_1 + M_4$, $Q_2 = M_1 + M_3$, $Q_3 = M_2$, and $Q_4 = M_3 - M_1$, we conclude that $[T]_{\mathcal{C}}$ is a diagonal matrix.



Example 37.6. We have shown that every linear transformation from a finite dimensional vector space V to V can be represented as a matrix transformation. As a consequence, all such linear transformations have eigenvalues. In this example we consider the problem of determining eigenvalues and eigenvectors of the linear transformation $T : V \rightarrow V$ defined by $T(f)(x) = \int_0^x f(t) dt$, where V is the vector space of all infinitely differentiable functions from \mathbb{R} to \mathbb{R} . Suppose that f is an eigenvector of T with a nonzero eigenvalue.

- (a) Use the Fundamental Theorem of Calculus to show that f must satisfy the equation $f'(x) = \frac{1}{\lambda}f(x)$ for some scalar nonzero scalar λ .
- (b) From calculus, we know that the functions f that satisfy the equation $f(x) = \frac{1}{\lambda}f'(x)$ all have the form $f(x) = Ae^{x/\lambda}$ for some scalar A . So if f is an eigenvector of T , then $f(x) = Ae^{x/\lambda}$ for some scalar A . Show, however, that $Ae^{x/\lambda}$ cannot be an eigenvector of T . Thus, T has no eigenvectors with nonzero eigenvalues.
- (c) Now show that 0 is not an eigenvalue of T . Conclude that T has no eigenvalues or eigenvectors.
- (d) Explain why this example does not contradict the statement that every linear transformation from a finite dimensional vector space V to V has eigenvalues.

Example Solution.

- (a) Assuming that f is an eigenvector of T with nonzero eigenvalue λ , then

$$\lambda f(x) = T(f)(x) = \int_0^x f(t) dt. \quad (37.2)$$

Recall that $\frac{d}{dx} \int_0^x f(t) dt = f(x)$ by the Fundamental Theorem of Calculus. Differentiating both sides of (37.2) with respect to x leaves us with $\lambda f'(x) = f(x)$, or $f'(x) = \frac{1}{\lambda}f(x)$.

- (b) We can directly check from the definition that $T(f)$ is not a multiple of $f(x) = Ae^{x/\lambda}$ unless $A = 0$, which is not allowed. Another method is to note that $T(f)(0) = \int_0^0 f(t) dt = 0$ by definition. But if $f(x) = Ae^{x/\lambda}$, then

$$0 = T(f)(0) = \lambda f(0) = Ae^{0/\lambda} = A.$$

This means that $f(x) = 0e^{x/\lambda} = 0$. But 0 can never be an eigenvector by definition. So T has no eigenvectors with nonzero eigenvalues.

- (c) Suppose that 0 is an eigenvalue of T . Then there is a nonzero function g such that $T(g) = 0$. In other words, $0 = \int_0^x g(t) dt$. Again, differentiating with respect to x yields the equation $g(x) = 0$. So T has no eigenvectors with eigenvalue 0. Since T does not have a nonzero eigenvalue or zero as an eigenvalue, T has no eigenvalues (and eigenvectors).
- (d) The reason this example does not contradict the statement is that V is an infinite dimensional vector space. In fact, the linearly independent monomials t^m are all in V for any positive integer m .



Summary

- We can define eigenvalues and eigenvectors of a linear transformation $T : V \rightarrow V$, where V is a finite dimensional vector space. In this case, a scalar λ is an eigenvalue for T if there exists a non-zero vector \mathbf{v} in V so that $T(\mathbf{v}) = \lambda\mathbf{v}$.
- To find eigenvalues and eigenvectors of a linear transformation $T : V \rightarrow V$, where V is a finite dimensional vector space, we find the eigenvalues and eigenvectors for $[T]_{\mathcal{B}}$, where \mathcal{B} is a basis for V . If \mathcal{C} is any other basis for V , then $[T]_{\mathcal{B}}$ and $[T]_{\mathcal{C}}$ are similar matrices and have the same eigenvalues. Once we find an eigenvector $[\mathbf{v}]_{\mathcal{B}}$ for $[T]_{\mathcal{B}}$, then \mathbf{v} is an eigenvector for T .
- A linear transformation $T : V \rightarrow V$, where V is a finite dimensional vector space, is diagonalizable if there is a basis \mathcal{C} for V for which $[T]_{\mathcal{C}}$ is a diagonalizable matrix.
- To determine if a linear transformation $T : V \rightarrow V$ is diagonalizable, we pick a basis \mathcal{C} for V . If the matrix $[T]_{\mathcal{C}}$ has n linearly independent eigenvectors, then T is diagonalizable.

Exercises

- (1) Let $V = \mathbb{P}_2$ and define $T : V \rightarrow V$ by $T(p(t)) = \frac{d}{dt}((1-t)p(t))$.
 - Show that T is a linear transformation.
 - Let \mathcal{S} be the standard basis for \mathbb{P}_2 . Find $[T]_{\mathcal{S}}$.
 - Find the eigenvalues and a basis for each eigenspace of T .
 - Is T diagonalizable? If so, find a basis \mathcal{B} for \mathbb{P}_2 so that $[T]_{\mathcal{B}}$ is a diagonal matrix. If not, explain why not.
- (2) Let $T : V \rightarrow V$ be a linear transformation, and let \mathcal{B} be a basis for V . Show that every eigenvector \mathbf{v} of T with eigenvalue λ corresponds to an eigenvector of $[T]_{\mathcal{B}}$ with eigenvalue λ .
- (3) Let C^∞ be the set of all functions from \mathbb{R} to \mathbb{R} that have derivatives of all orders.
 - Explain why C^∞ is a subspace of \mathcal{F} .
 - Let $D : C^\infty \rightarrow C^\infty$ be defined by $D(f) = f'$. Explain why D is a linear transformation.
 - Let λ be any real number and let f_λ be the exponential function defined by $f_\lambda(x) = e^{\lambda x}$. Show that f_λ is an eigenvector of D . What is the corresponding eigenvalue? How many eigenvalues does D have?
- (4) Consider $D : \mathbb{P}_2 \rightarrow \mathbb{P}_2$, where D is the derivative operator defined as in Exercise 3.. Find all of the eigenvalues of D and a basis for each eigenspace of D .
- (5) Let n be a positive integer and define $T : \mathcal{M}_{n \times n} \rightarrow M_{n \times n}$ by $T(A) = A^T$.
 - Show that T is a linear transformation.



- (b) Is $\lambda = 1$ an eigenvalue of T ? Explain. If so, describe in detail the vectors in the corresponding eigenspace.
- (c) Does T have any other eigenvalues? If so, what are they and what are the vectors in the corresponding eigenspaces? If not, why not?
- (6) Label each of the following statements as True or False. Provide justification for your response.
- True/False** The number 0 cannot be an eigenvalue of a linear transformation.
 - True/False** The zero vector cannot be an eigenvector of a linear transformation.
 - True/False** If \mathbf{v} is an eigenvector of a linear transformation T , then so is $2\mathbf{v}$.
 - True/False** If \mathbf{v} is an eigenvector of a linear transformation T , then \mathbf{v} is also an eigenvector of the transformation $T^2 = T \circ T$.
 - True/False** If \mathbf{v} and \mathbf{u} are eigenvectors of a linear transformation T with the same eigenvalue, then $\mathbf{v} + \mathbf{u}$ is also an eigenvector of T with the same eigenvalue.
 - True/False** If λ is an eigenvalue of a linear transformation T , then λ^2 is an eigenvalue of T^2 .
 - True/False** Let S and T be two linear transformations with the same domain and codomain. If \mathbf{v} is an eigenvector of both S and T , then \mathbf{v} is an eigenvector of $S + T$.
 - True/False** Let V be a vector space and let $T : V \rightarrow V$ be a linear transformation. Then T has 0 as an eigenvalue if and only if $[T]_{\mathcal{B}}$ has linearly dependent columns for any basis \mathcal{B} of V .

Project: Linear Transformations and Differential Equations

There are many different types of differential equations, but we will focus on differential equations of the form $my'' = -ky - by'$ presented in the introduction, a second order (the highest derivative in the equation is a second order derivative) linear (the coefficients are constants) differential equation (also called damped harmonic oscillators).

To solve a differential equation means to find all solutions to the differential equation. That is, find all functions y that satisfy the differential equation. For example, since $\frac{d}{dt}t^2 = 2t$, we see that $y = t^2$ satisfies the differential equation $y' = 2t$. But t^2 is not the only solution to this differential equation. In fact, $y = t^2 + C$ is a solution to $y' = 2t$ for any scalar C . We will see how to represent solutions to the differential equation $my'' = -ky - by'$ in this project.

The next activity shows that the set of solutions to the linear differential equation $my'' = -ky - by'$ is a subspace of the vector space \mathcal{F} of all functions from \mathbb{R} to \mathbb{R} . So we should expect close connections between differential equations and linear algebra, and we will make some of these connections as we proceed.

Project Activity 37.1. We can represent differential equations using linear transformations. To see how, let D be the function from the space C^1 of all differentiable real-valued functions to \mathcal{F} given by $D(f) = \frac{df}{dt}$.



- (a) Show that D is a linear transformation.
- (b) In order for a function f to be a solution to a differential equation of the form (37.1), it is necessary for f to be twice differentiable. We will assume that D acts only on such functions from this point on. Use the fact that D is a linear transformation to show that the differential equation $my'' = -ky - by'$ can be written in the form

$$(mD^2 + bD)(y) = -ky.$$

Project Activity 37.1 shows that any solution to the differential equation $my'' = -ky - by'$ is an eigenvector for the linear transformation $mD^2 + bD$. That is, the solutions to the differential equation $my'' = -ky - by'$ form the eigenspace E_{-k} of $mD^2 + bD$ with eigenvalue $-k$. The eigenvectors for a linear transformation acting on a function space are also called *eigenfunctions*. To build up to solutions to the second order differential equation we have been considering, we start with solutions to the first order equation.

Project Activity 37.2. Let k be a scalar. Show that the solutions of the differential equation $D(y) = -ky$ form a one-dimensional subspace of \mathcal{F} . Find a basis for this subspace. Note that this is the eigenspace of the transformation D corresponding to the eigenvalue $-k$. (Hint: To find the solutions to $y' = -ky$, write y' as $\frac{dy}{dt}$ and express the equation $y' = -ky$ in the form $\frac{dy}{dt} = -ky$. Divide by y and integrate with respect to t .)

Before considering the general second order differential equation, we start with a simpler example.

Project Activity 37.3. As a specific example of a second order linear equation, as discussed at the beginning of this section, Hooke's law states that if a mass is hanging from a spring, the force acting on the spring is proportional to the displacement of the spring from equilibrium. If we let y be the displacement of the object from its equilibrium, and ignore any resistance, the position of the mass-spring system can be represented by the differential equation

$$mD^2(y) = -ky,$$

where m is the mass of the object and k is a positive constant that depends on the spring. Assuming that the mass is positive, we can divide both sides by m and rewrite this differential equation in the form

$$D^2(y) = -cy \tag{37.3}$$

where $c = \frac{k}{m}$. So the solutions to the differential equation (37.3) make up the eigenspace E_{-c} for D^2 with eigenvalue $-c$.

- (a) Since the derivatives of $\sin(t)$ and $\cos(t)$ are scalar multiples of $\sin(t)$ and $\cos(t)$, it may be reasonable that these make up solutions to (37.3). Show that $y_1 = c\cos(t)$ and $y_2 = c\sin(t)$ are both in E_{-c} .
- (b) As functions, the cosine and sine are related in many ways (e.g., the Pythagorean Identity). An important property for this application is the linear independence of the cosine and sine. Show, using the **definition** of linear independence, that the cosine and sine functions are linearly independent in \mathcal{F} .



- (c) Part (a) shows that there are at least two different functions in E_{-c} . To solve a differential equation is to find all of the solutions to the differential equation. In other words, we want to completely determine the eigenspace E_{-c} . We have already seen that any function y of the form $y(t) = c_1 \cos(t) + c_2 \sin(t)$ is a solution to the differential equation $mD^2(y) = -ky$. The theory of linear differential equations tells us that there is a unique solution to $mD^2(y) = -ky$ if we specify two initial conditions. What this means is that to show that any solution z to the differential equation $mD^2(y) = -ky$ with two initial values $z(t_0)$ and $z'(t_0)$ for some scalar t_0 is of the form $y = c_1 \cos(t) + c_2 \sin(t)$, we need to verify that there are values of c_1 and c_2 such that $y(t_0) = z(t_0)$ and $y'(t_0) = z'(t_0)$. Here we will use this idea to show that any function in E_{-c} is a linear combination of $\cos(t)$ and $\sin(t)$. That is, that the set $\{\cos, \sin\}$ spans E_{-c} . Let $y = c_1 \cos(t) + c_2 \sin(t)$. Show that there are values for c_1 and c_2 such that

$$y(0) = z'(0) \text{ and } y'(0) = z'(0).$$

This result, along with part(b), shows that $\{\cos, \sin\}$ is a basis for E_{-c} . (Note: That the solutions to differential equation (37.3) involve sines and cosines models the situation that a mass hanging from a spring will oscillate up and down.)

As we saw in Project Activity 37.2, the eigenspace E_k of the linear transformation D is one-dimensional. The key idea in Project Activity 37.3 that allowed us to find a basis for the eigenspace of D^2 with eigenvalue $-c$ is that \cos and \sin are linearly independent eigenfunctions that span E_{-c} . We won't prove this result, but the general theory of linear differential equations states that if y_1, y_2, \dots, y_n are linearly independent solutions to the n th order linear differential equation

$$(a_n D^n + a_{n-1} D^{n-1} + \cdots + a_1 D)(y) = -a_0 y,$$

then $\{y_1, y_2, \dots, y_n\}$ is a basis for the eigenspace of the linear transformation $a_n D^n + a_{n-1} D^{n-1} + \cdots + a_1 D$ with eigenvalue $-a_0$. Any basis for the solution set to a differential equation is called a *fundamental set of solutions* for the differential equation. Consequently, it is important to be able to determine when a set of functions is linearly independent. One tool for doing so is the Wronskian, which we study in the next activity.

Project Activity 37.4. Suppose we have n functions f_1, f_2, \dots, f_n , each with $n - 1$ derivatives. To determine the independence of the functions we must understand the solutions to the equation

$$c_1 f_1(t) + c_2 f_2(t) + \cdots + c_n f_n(t) = 0. \quad (37.4)$$

We can differentiate both sides of Equation (37.4) to obtain the new equation

$$c_1 f'_1(t) + c_2 f'_2(t) + \cdots + c_n f'_n(t) = 0.$$

We can continue to differentiate as long as the functions are differentiable to obtain the system

$$\begin{aligned} f_1(t)c_1 + & f_2(t)c_2 + \cdots + f_n(t)c_n = 0 \\ f'_1(t)c_1 + & f'_2(t)c_2 + \cdots + f'_n(t)c_n = 0 \\ f''_1(t)c_1 + & f''_2(t)c_2 + \cdots + f''_n(t)c_n = 0 \\ & \vdots \\ f_1^{(n-1)}(t)c_1 + & f_2^{(n-1)}(t)c_2 + \cdots + f_n^{(n-1)}(t)c_n = 0. \end{aligned}$$

- (a) Write this system in matrix form, with coefficient matrix

$$A = \begin{bmatrix} f_1(t) & f_2(t) & \cdots & f_n(t) \\ f'_1(t) & f'_2(t) & \cdots & f'_n(t) \\ \vdots & & & \\ f_1^{(n-1)}(t) & f_2^{(n-1)}(t) & \cdots & f_n^{(n-1)}(t) \end{bmatrix}.$$

- (b) The matrix in part (a) is called the Wronskian matrix of the system. The scalar

$$W(f_1, f_2, \dots, f_n) = \det(A)$$

is called the Wronskian of f_1, f_2, \dots, f_n . What must be true about the Wronskian for our system to have a unique solution? If the system has a unique solution, what is the solution? What does this result tell us about the functions f_1, f_2, \dots, f_n ?

- (c) Use the Wronskian to show that the cosine and sine functions are linearly independent.

We can apply the Wronskian to help find bases for the eigenspace of the linear transformation $mD^2 + bD$ with eigenvalue k .

Project Activity 37.5. The solution to the Hooke's Law differential equation in Project Activity 37.3 indicates that the spring will continue to oscillate forever. In reality, we know that this does not happen. In the non-ideal case, there is always some force (e.g., friction, air resistance, a physical damper as in a piston) that acts to dampen the motion of the spring causing the oscillations to die off. Damping acts to oppose the motion, and we generally assume that the faster an object moves, the higher the damping. For this reason we assume the damping force is proportional to the velocity. That is, the damping force has the form $-by'$ for some positive constant b . This produces the differential equation

$$my'' + by' + ky = 0 \quad (37.5)$$

or $(mD^2 + bD) = -ky$. We will find bases for the eigenspace of the linear transformation $mD^2 + bD$ with eigenvalue $-k$ in this activity.

- (a) Since derivatives of exponential functions are still exponential functions, it seems reasonable to try an exponential function as a solution to (37.5). Show that if $y = e^{rt}$ for some constant r is a solution to (37.5), then $mr^2 + br + k = 0$. The equation $mr^2 + br + k = 0$ is the characteristic or auxiliary equation for the differential equation.
- (b) Part (a) shows that our solutions to the differential equation (37.5) are exponential of the form e^{rt} , where r is a solution to the auxiliary equation. Recall that if we can find two linearly independent solutions to (37.5), then we have found a basis for the eigenspace E_{-k} of $mD^2 + bD$ with eigenvalue $-k$. The quadratic equation shows that the roots of the auxiliary equation are

$$\frac{-b \pm \sqrt{b^2 - 4mk}}{2m}.$$

As we will see, our basis depends on the types of roots the auxiliary equation has.



- i. Assume that the roots r_1 and r_2 of the auxiliary equation are real and distinct. That means that $y_1 = e^{r_1 t}$ and $y_2 = e^{r_2 t}$ are eigenfunctions in E_{-k} . Use the Wronskian to show that $\{y_1, y_2\}$ is a basis for E_{-k} in this case. Then describe the behavior of an arbitrary eigenfunction in E_{-2} if $mD^2 + bD = D^2 + 3D$ and how it relates to damping. Draw a representative solution to illustrate. (In this case we say that the system is *overdamped*. These systems can oscillate at most once, then they quickly damp to 0.)
- ii. Now suppose that we have a repeated real root r to the auxiliary equation. Then there is only one exponential function $y_1 = e^{rt}$ in E_{-k} . In this case, show that $y_2 = te^{rt}$ is also in E_{-k} and that $\{y_1, y_2\}$ is a basis for E_{-k} . Then describe the behavior of an arbitrary eigenfunction in E_{-1} if $mD^2 + bD = D^2 + 2D$ and how it relates to damping. Draw a representative solution to illustrate. (In this case we say that the system is *critically damped*. These systems behave similar to the overdamped systems in that they do not oscillate. However, if the damping is reduced just a little, the system can oscillate.)
- iii. The last case is when the auxiliary equation has complex roots $z_1 = u + vi$ and $z_2 = u - vi$. We want to work with real valued functions, so we need to determine real valued solutions from these complex roots. To resolve this problem, we note that if x is a real number, then $e^{ix} = \cos(x) + i \sin(x)$. So

$$e^{(u+vi)t} = e^{ut}e^{ivt} = e^{ut} \cos(vt) + e^{ut} \sin(vt)i.$$

Show that $\{e^{ut} \cos(vt), e^{ut} \sin(vt)\}$ is a basis for E_{-k} in this case. Then describe the behavior of an arbitrary eigenfunction in E_{-5} if $mD^2 + bD = D^2 + 2D$ and how it relates to damping. Draw a representative solution to illustrate. (In this case we say that the system is *underdamped*. These systems typically exhibit some oscillation.)

Project Activity 37.4 tells us that if $W(f_1, f_2, \dots, f_n)$ is not zero, then f_1, f_2, \dots, f_n are linearly independent. You might wonder what conclusion we can draw if $W(f_1, f_2, \dots, f_n)$ is zero.

Project Activity 37.6. In this activity we consider the Wronskian of two different pairs of functions.

(a) Calculate $W(t, 2t)$. Are t and $2t$ linearly independent or dependent? Explain.

(b) Now let $f(t) = t|t|$ and $g(t) = t^2$.

- i. Calculate $f'(t)$ and $g'(t)$. (Hint: Recall that $|x| = \begin{cases} x & \text{if } x \geq 0 \\ -x & \text{if } x < 0. \end{cases}$)
- ii. Calculate $W(f, g)$. Are f and g linearly independent or dependent in \mathcal{F} ? Explain. (Hint: Consider the cases when $t \geq 0$ and $t < 0$.)
- iii. What conclusion can we draw about the functions f_1, f_2, \dots, f_n if $W(f_1, f_2, \dots, f_n)$ is zero? Explain.

Appendix A

Complex Numbers

Focus Questions

By the end of this section, you should be able to give precise and thorough answers to the questions listed below. You may want to keep these questions in mind to focus your thoughts as you complete the section.

- What is a complex number?
- How is the sum and product of two complex numbers defined?
- How do we find the multiplicative inverse of a nonzero complex number?
- What general structure does the set of complex numbers have?

Complex Numbers

Complex numbers are usually introduced as a tool to solve the quadratic equation $x^2 + 1 = 0$. However, that is not how complex numbers first came to light. The story actually involves solutions to the general cubic equation. The interested reader could consult Chapter 6 of William Dunham's excellent book *Journey Through Genius*. In this appendix we touch on the basics of complex numbers to provide enough context for the section on complex eigenvalues.

A complex number is defined by a pair of real numbers - the *real part* of the complex number and the *imaginary part* of the complex number.

Definition A.1. A **complex number** is a number of the form

$$a + bi$$

where a and b are real numbers and $i^2 = -1$.

The number a is the real part of the complex number and the number b is the imaginary part. We often write

$$z = a + bi$$

for a complex number z . We say that two complex numbers $a + bi$ and $c + di$ are equal if $a = c$ and $b = d$.

There is an arithmetic of complex numbers that is determined by an addition and multiplication of complex numbers. Adding complex numbers is natural:

$$(a + bi) + (c + di) = (a + c) + (b + d)i.$$

That is, to add two complex numbers we add their real parts together and their imaginary parts together.

Activity A.1. Multiplication of complex numbers is also done in a natural way.

- (a) By expanding the product as usual, treating i as we would any real number, and exploiting the fact that $i^2 = -1$, explain why we define the product of complex numbers $a + bi$ and $c + di$ as

$$(a + bi)(c + di) = (ac - bd) + (bc + ad)i.$$

- (b) Use the definitions of addition and multiplication to write each of the sums or products as a complex number in the form $a + bi$.

- i. $(2 + 3i) + (7 - 4i)$
- ii. $(4 - 2i)(3 + i)$
- iii. $(2 + i)i - (3 + 4i)$

It isn't difficult to show that the set of complex numbers, which we denote by \mathbb{C} , satisfies many useful and familiar properties.

Activity A.2. Show that \mathbb{C} has the same structure as \mathbb{R} . That is, show that for all u, w , and z in \mathbb{C} , the following properties are satisfied.

- (a) $w + z \in \mathbb{C}$ and $wz \in \mathbb{C}$
- (b) $w + z = z + w$ and $wz = zw$
- (c) $(w + z) + u = w + (z + u)$ and $(wz)u = w(zu)$
- (d) There is an element 0 in \mathbb{C} such that $z + 0 = z$
- (e) There is an element 1 in \mathbb{C} such that $(1)z = z$
- (f) There is an element $-z$ in \mathbb{C} such that $z + (-z) = 0$
- (g) If $z \neq 0$, there is an element $\frac{1}{z}$ in \mathbb{C} such that $z\left(\frac{1}{z}\right) = 1$
- (h) $u(w + z) = (uw) + (uz)$

The result of Activity A.2 is that, just like \mathbb{R} , the set \mathbb{C} is a field. If we wanted to, we could define vector spaces over \mathbb{C} just like we did over \mathbb{R} . The same results hold.



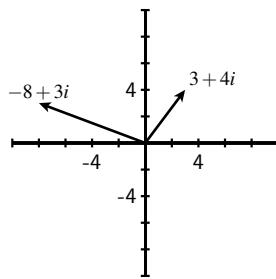


Figure A.1: Two complex numbers.

Conjugates and Modulus

We can draw pictures of complex numbers in the plane. We let the x -axis be the real axis for a complex number and the y -axis the imaginary axis. That is, if $z = a + bi$ we can think of z as a directed line segment from the origin to the point (a, b) , where the terminal point of the segment is a units from the imaginary axis and b units from the real axis. For example, the complex numbers $3 + 4i$ and $-8 + 3i$ are shown in Figure A.1.

We can also think of the complex number $z = a + bi$ as the vector $[a \ b]^T$. In this way, the set \mathbb{C} is a two-dimensional vector space over \mathbb{R} with basis $\{1, i\}$. Each of these complex numbers has a length that we call the *norm* or *modulus* of the complex number. We denote the norm of a complex number $a + bi$ as $|a + bi|$. The distance formula or the Pythagorean theorem show that

$$|a + bi| = \sqrt{a^2 + b^2}.$$

Note that

$$a^2 + b^2 = a^2 - b^2i^2 = (a + bi)(a - bi)$$

so the norm of the complex number $a + bi$ can also be viewed as a square root of the product of $a + bi$ with $a - bi$. The number $a - bi$ is called the *complex conjugate* of $a + bi$. If we let $z = a + bi$, we denote the complex conjugate of z as \bar{z} . So $\overline{a + bi} = a - bi$.

Activity A.3. Let $w = 2 + 3i$ and $z = -1 + 5i$.

- (a) Find \bar{w} and \bar{z} .
- (b) Compute $|w|$ and $|z|$.
- (c) Compute $w\bar{w}$ and $z\bar{z}$.
- (d) Let z be an arbitrary complex number. There is a relationship between $|z|$, z , and \bar{z} . Find and verify this relationship.
- (e) What is \bar{z} if $z \in \mathbb{R}$?

Appendix B

Answers and Hints to Selected Exercises

Section 1

1. (a) Consider the equations

$$b_1x_1 + b_2x_2 + b_3x_3 = s,$$

and

$$k(b_1x_1 + b_2x_2 + b_3x_3) = ks.$$

- (b) Similar to part (a)

3. A cup of coffee costs \$1.95, a muffin costs \$2.05, and a scone costs \$2.15.

5. (a)

$$x_1 = \frac{x_2 + 200 + 0 + 0}{4}$$

$$x_2 = \frac{x_1 + x_3 + 200 + 0}{4}$$

$$x_3 = \frac{x_2 + 0 + 400 + 0}{4}$$

- (b) $x_1 = 75$, $x_2 = 100$ and $x_3 = 125$.

7. (a)

$$x_1 + x_2 + x_3 = 0$$

$$x_1 + x_2 + x_3 = 1.$$

- (b) Impossible.

- (c) Impossible.

(d)

$$x_1 + x_2 + x_3 = 0$$

$$2x_1 + 2x_2 + 2x_3 = 0.$$

9. (a) This system has the solution $x_1 = \frac{h-10}{3h-5}$ and $x_2 = \frac{5}{3h-5}$ as long as $h \neq \frac{5}{3}$.

- (b) When $h = \frac{5}{3}$, $x_1 = 2 - \frac{5}{3}x_2$ and x_2 is free.

11. (a) F

- (c) F

- (e) F

Section 2

1. (a) $h = 6$

- (b) $h = 6$ and $k \neq -2$.

- (c) $h = 6$ and $k = -2$.

3.

$$x_1 + x_2 = 1$$

$$-x_1 + x_2 = -3$$

$$x_1 + x_2 + x_3 = 1.$$

5. Student 2's hunch is correct.

6. (a) F

- (c) T

- (e) T

(g) F

(i) F

(k) F

Section 3

1. Augmented matrix:

$$\left[\begin{array}{ccc|c} 1 & 1 & -1 & 4 \\ 1 & 2 & 2 & 3 \\ 2 & 3 & -3 & 11 \end{array} \right],$$

solution $x_1 = 1$, $x_2 = 2$, and $x_3 = -1$

3. Row operations show this.

- 5.
- $x_1 = -4 - 6x_3$
- ,
- $x_2 = -3 - 4x_3$
- ,
- x_3
- is free,
- $x_4 = 3$

7. (a) Yes

$$(b) \left[\begin{array}{cc|c} \blacksquare & * & * \\ 0 & \blacksquare & * \end{array} \right], \left[\begin{array}{cc|c} 0 & \blacksquare & * \\ 0 & 0 & * \end{array} \right]$$

9. This is not possible.

$$11. \left[\begin{array}{cc} \blacksquare & * \\ 0 & * \end{array} \right] \text{ and } \left[\begin{array}{cc} 0 & * \\ 0 & 0 \end{array} \right].$$

- 12 (a) T

(c) T

(e) F

(g) T

(i) F

(k) F

Section 4

- 1.
- $\mathbf{w} = \left(-\frac{9}{4}\right)\mathbf{u} + \left(\frac{7}{4}\right)\mathbf{v}$

3. Only when
- $w_1 = 2w_2 - 3w_3$
- .
- $0\mathbf{u} + 0\mathbf{v} = \mathbf{0}$

5. (a) We cannot make the desired solution.

- (b) We can make the desired chemical solution with
- $\frac{3}{4}$
- of solution
- \mathbf{v}_1
- for every
- $\frac{1}{4}$
- of solution
- \mathbf{v}_2
- .

(c) If $\mathbf{w} = \begin{bmatrix} w_1 \\ w_2 \\ w_3 \end{bmatrix}$ can be made from our original two chemical solutions, then $w_3 + \frac{3}{23}w_1 - \frac{7}{23}w_2 = 0$.

7. (a)
- $\text{Span}\{\mathbf{v}_1\}$
- is the line in
- \mathbb{R}^2
- through the origin and the point
- $(1, 1)$
- .

- (b)
- $\text{Span}\{\mathbf{v}_1, \mathbf{v}_3\}$
- is the plane in
- \mathbb{R}^3
- through the origin and the points
- $(1, 1, 1)$
- and
- $(2, 0, 1)$
- .

9. Yes.

11. (a) F

(c) T

(e) T

(g) F

(i) T

(k) T

(m) F

Section 5

1. The matrix-vector form of the system is

$$A\mathbf{x} = \mathbf{b}, \text{ where } \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}.$$

3. The system corresponding to this matrix-vector equation is

$$2x_1 + 3x_2 + 4x_3 = 4$$

$$x_1 - 2x_2 + 3x_3 = -6.$$

The solution to the system is

$$x_1 = -\frac{17}{7}x_3 - \frac{10}{7}$$

$$x_2 = \frac{2}{7}x_3 + \frac{16}{7}$$

 x_3 is free.

5. Both methods give

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_1a + x_2b \\ x_1c + x_2d \end{bmatrix}.$$



Section 7

7. (a) b_3 (b) If $b_3 = 9$, then $a = 2$.

$$9. A \begin{bmatrix} -1 \\ 6 \\ -5 \end{bmatrix} = \begin{bmatrix} 3 \\ -5 \end{bmatrix}.$$

11. The general solution to the system is

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = y \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix} + z \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} + \begin{bmatrix} 3 \\ 0 \\ 0 \end{bmatrix}.$$

The solution is the plane in \mathbb{R}^3 through the points $(3, 0, 0)$, $(5, 1, 0)$, and $(2, 0, 1)$.

$$13. \mathbf{x} = \begin{bmatrix} 2 \\ -1 \\ -1 \end{bmatrix}$$

14. (a) F

(c) T

(e) T

(g) T

(i) T

(k) T

$$1. \text{ If } \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}, \text{ then } T(\mathbf{x}) = \begin{bmatrix} x_1 \\ x_1 + 2x_2 + x_3 \\ x_1 - 3x_3 \end{bmatrix}.$$

$$3. \begin{bmatrix} 12 \\ -11 \end{bmatrix}.$$

$$5. A = \begin{bmatrix} 2 & -1 \\ 1 & 3 \end{bmatrix}$$

$$7. T \text{ defined by } T(\mathbf{x}) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \mathbf{x}$$

9. Not possible.

$$11. L = \mathbf{v} + c\mathbf{w}, T(L) = T(\mathbf{v} + c\mathbf{w}) = T(\mathbf{v}) + cT(\mathbf{w})$$

12. (a) F

(c) F

(e) T

(g) F

(i) F

(k) F

(m) T

(o) T

Section 6

1. Linearly dependent. Change the last entry in \mathbf{v}_3 to a 0.

3. Only two of the solutions are necessary.

5. (a) $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$

$$(b) \mathbf{u} = \mathbf{v}_1 - 2\mathbf{v}_2 + 3\mathbf{v}_3$$

7. This is not possible.

9. (a) F

(c) T

(e) T

(g) T

(i) T

(k) T

Section 8

$$1. (a) AB = \begin{bmatrix} a & b \\ c & d \\ 0 & 0 \end{bmatrix}$$

$$(b) AB = [-2 \quad -2].$$

$$3. (a) X = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$$

(b) There is no matrix X with this property.



(c) $X = \begin{bmatrix} 2x_{21} & 1+2x_{22} \\ x_{21} & x_{22} \end{bmatrix}$,
where x_{21} and x_{22} can be any
scalars

5. $A^m \mathbf{v} = 2^m \mathbf{v}$

7. $A^2 = 0$ and $B^3 = 0$

9. (a) $[a_{ij} + b_{ij}] = [b_{ij} + a_{ij}]$

(c) $[a_{ij} + 0] = [a_{ij}]$

(e) $[(a+b)a_{ij}] = [aa_{ij}] + [ba_{ij}]$

(g) $[(ab)a_{ij}] = a[ba_{ij}]$

11. (a) Let $A^\top = [a'_{ij}]$. Then $a'_{ij} = a_{ji}$ by definition of the transpose. Let $(A^\top)^\top = [a''_{ij}]$. Then $a''_{ij} = a'_{ji} = a_{ij}$. So the ij th entry of $(A^\top)^\top$ is the same as the ij th entry of A , and we conclude that $(A^\top)^\top = A$.

(c) The ij th entry of aA is aa_{ij} , so the ij th entry of $(aA)^\top$ is aa_{ji} . But aa_{ji} is also the ij th entry of aA^\top . We conclude that $(aA)^\top = aA^\top$.

13. If $A = \begin{bmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{bmatrix}$, $B = \begin{bmatrix} \cos(\beta) & -\sin(\beta) \\ \sin(\beta) & \cos(\beta) \end{bmatrix}$, then AB equals $\begin{bmatrix} \cos(\alpha + \beta) & -\sin(\alpha + \beta) \\ \sin(\alpha + \beta) & \cos(\alpha + \beta) \end{bmatrix}$

14. (a) F

(c) F

(e) F

(g) T

Section 9

1. (a) Eigenvector with eigenvalue -1 .

(b) Eigenvector with eigenvalue 3 .

(c) Eigenvector with eigenvalue 2 .

3. (a) Eigenvalue

(b) Eigenvalue

(c) Not an eigenvalue

(d) Not an eigenvalue

5. (a) $a_3 = 804,000$ $s_3 = 196,000$,
 $a_4 = 782,400$, $s_4 = 217,600$

(b) $a_{k+1} = 0.9a_k + 0.3s_k$, $s_{k+1} = 0.1a_k + 0.7s_k$

(c) $\begin{bmatrix} a_{k+1} \\ s_{k+1} \end{bmatrix}$ equals
 $\begin{bmatrix} 0.9 & 0.3 \\ 0.1 & 0.7 \end{bmatrix} \begin{bmatrix} a_k \\ s_k \end{bmatrix}$

6. (a) F

(c) T

(e) T

(g) T

(i) F

(k) T

Section 10

1. $C = (I)C = C = (BA)C = BAC = B(AC) = B(I) = B$

3. (a) $\begin{bmatrix} 1 & k \\ 0 & 1 \end{bmatrix}^{-1} = \begin{bmatrix} 1 & -k \\ 0 & 1 \end{bmatrix}$

(b) Row reduce $[A | I_2]$

(c) Row reduce $[A | I_3]$ to see that

$$A^{-1} = \begin{bmatrix} 1 & -k & mk - \ell \\ 0 & 1 & -m \\ 0 & 0 & 1 \end{bmatrix}.$$

5. $c \neq -4$

8. (a) T

(c) T

(e) T

(g) T

(i) F

Section 11



1. A is invertible as long as $\begin{bmatrix} a \\ b \\ c \end{bmatrix}$ is not in the plane in \mathbb{R}^3 through the origin and the points $(1, -1, 1)$ and $(2, 1, 1)$
3. (a) T
(c) T
(e) T
(g) T
11. Not in general a subspace of \mathbb{R}^n .
13. (a) F
(c) T
(e) T
(g) F
(i) T
(k) F

Section 13

Section 12

1. (a) Closed under addition, not closed under multiplication by scalars, contains the zero vector.
(b) Not closed under addition, closed under multiplication by scalars, contains the zero vector.
(c) Not closed under addition, not closed under multiplication by scalars, does not contain the zero vector.
(d) Not closed under addition, closed under multiplication by scalars, contains the zero vector.
3. The only such set is the empty set.
5. \mathbb{R}^2 .

7. (a) As an example, let $\mathbf{v} = [2 \ 1]^\top$ in \mathbb{R}^2 .

- i. I_2
- ii. $2I_2$
- iii. $[3\mathbf{e}_1 - \mathbf{e}_2]$
- iv. $\left[\frac{a}{2}\mathbf{e}_1 \ b\mathbf{e}_2\right] [2 \ 1]^\top = a\mathbf{e}_1 + b\mathbf{e}_2 = [a \ b]^\top$

- (b) $\mathbf{w}_1 + \mathbf{w}_2 = A_1\mathbf{v} + A_2\mathbf{v} = (A_1 + A_2)\mathbf{v}$, $c\mathbf{w}_1 = c(A_1\mathbf{v}) = A_1(c\mathbf{v})$, $0\mathbf{v} = \mathbf{0}$.

- (c) $\{\mathbf{0}\}, \mathbb{R}^m$

9. $W_1 \cap W_2$ is a subspace of \mathbb{R}^n , $W_1 \cup W_2$ is not in general a subspace of \mathbb{R}^n

1. A basis for $\text{Col } A$ is $\left\{ \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 3 \\ 2 \\ 5 \end{bmatrix} \right\}$, a basis for $\text{Nul } A$ is $\left\{ \begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} -7 \\ 0 \\ 1 \\ 1 \end{bmatrix} \right\}$.

3. $a = -3$ and $b = 0$
5. $\begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix}$
7. Not possible
8. (a) T
(c) T
(e) T
(g) F
(i) F

Section 14

1. (a) $\left\{ \begin{bmatrix} -1 \\ 1 \end{bmatrix} \right\}$
(b) $\left\{ \begin{bmatrix} -2 \\ 1 \end{bmatrix} \right\}$
(c) $\left\{ \begin{bmatrix} -1 \\ 1 \end{bmatrix} \right\}$
(d) $\left\{ \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} \right\}$



(e) $\left\{ \begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix} \right\}$

(f) $\left\{ \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -2 \\ 0 \\ 1 \end{bmatrix} \right\}$

3. $a = 0, b = 4$

5.

- (a) Eigenvalue 0, eigenspace \mathbb{R}^2
 (b) Eigenvalue 0, eigenspace \mathbb{R}^n

6. (a) F

(c) F

(e) T

(g) T

Section 15

1. (a) $\left\{ \begin{bmatrix} 1 \\ 0 \\ 2 \\ 1 \\ 3 \\ 3 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 0 \\ 2 \\ 1 \\ 3 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \\ -1 \\ 1 \end{bmatrix} \right\}$,
 $\dim(\text{Col } A) = 3, \dim(\text{Nul } A) = 2$

(b) $\left\{ \begin{bmatrix} -3 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ -2 \\ 1 \\ 0 \end{bmatrix} \right\}$

3. (a) $\left\{ \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix} \right\}$, $\text{rank}(A) = 1$

(b) $\left\{ \begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix} \right\}$,
 $\text{nullity}(A) = 3$

(c) $\text{rank}(A) + \text{nullity}(A) = 1 + 3 = 4$

(d) $\left\{ \begin{bmatrix} 1 \\ 2 \\ -1 \\ -1 \end{bmatrix} \right\}$, dimension 1

5. (a) What are the solutions to $A\mathbf{x} = \mathbf{0}$?

(b) Use part (a) and the Rank-Nullity Theorem.

7. $\dim(W)$ can be 0, 1, 2, 3, or 4 corresponding to $\{\mathbf{0}\}$, a line through the origin in \mathbb{R}^4 , a plane containing the origin in \mathbb{R}^4 , a copy of \mathbb{R}^3 through the origin in \mathbb{R}^4 , \mathbb{R}^4

9. $\dim(\text{Col } A) = 3, \dim(\text{Nul } A) = 2$

11.

(a) Not possible.

(b) $\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$

13. (a) F

(c) F

(e) T

(g) T

(i) F

(k) T

Section 16

1. $\det(2A) =$
 $(8a_{11})[(a_{22})(a_{33}) - (a_{23})(a_{32})] -$
 $(8a_{12})[(a_{21})(a_{33}) - (a_{23})(a_{31})] +$
 $(8a_{13})[(a_{21})(a_{32}) - (a_{22})(a_{31})] = 8 \det(A).$

3. (a) $\det(A^2) = \det(AA) = \det(A)\det(A) = [\det(A)]^2$.

(b) $\det(A^k) = \det(AA^{k-1}) = \det(A)[\det(A)]^{k-1} = [\det(A)]^k$.

(c) Yes.

5. $(\det(A))^2$

7. (a) F

(c) F

(e) T

- (g) F
(i) T
(k) F

Section 17

1. (a) i. $\det(B - \lambda I_2) = (1 - \lambda)(-2 - \lambda) - 2 = \lambda^2 + \lambda - 4$
ii. $B^2 + B - 4I_2 = 0$.

(b) $\det(A - AI_n) = \det(0) = 0$.

3. (a) Eigenvalues 0, 3; algebraic multiplicities 2, 1; geometric multiplicities 2,1
(b) Eigenvalues 1, 2; algebraic multiplicities 1, 2; geometric multiplicities 1,1

5. $\begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}, \begin{bmatrix} 3 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 2 \end{bmatrix}, \begin{bmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}$

7. (a) F
(c) T
(e) F
(g) T
(i) T

Section 18

1. (a) Not diagonalizable.

(b) Diagonalizable by $P = \begin{bmatrix} 1 & 2 & 1 \\ 1 & 3 & 3 \\ 1 & 3 & 4 \end{bmatrix}$

3. $A = \begin{bmatrix} 1 & 2 \\ 0 & 2 \end{bmatrix}, P_1 = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}, D_1 = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}, P_2 = \begin{bmatrix} 2 & 1 \\ 1 & 0 \end{bmatrix}, D_2 = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}$

5. Yes

7.

- (a) Eigenvalues, 1; eigenvectors : $\text{Span}\{[1 \ 0]^\top\}$

- (b) Diagonalizable

9. (a) The ii entry of RS is $r_{i1}s_{1i} + r_{i2}s_{2i} + \dots + r_{in}s_{ni}$. The jj entry of SR is $s_{j1}r_{1j} + s_{j2}r_{2j} + \dots + s_{jn}r_{nj}$. Sum as i and j go from 1 to n .

(b) i. $\text{trace}(D) = \text{trace}(P^{-1}(AP)) = \text{trace}((AP)P^{-1}) = \text{trace}(A)$
ii. $\text{trace}(A) = \text{trace}(D) = \sum_{i=1}^n \lambda_i$

11. (a) $e^A = \begin{bmatrix} e & e-1 \\ 0 & 1 \end{bmatrix}$

(b) $e^B = I_2 + B = \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix}$.

(c) $e^{A+B} = \begin{bmatrix} e & 0 \\ 0 & 1 \end{bmatrix}$

- (d) No

13. (a)

$$\begin{aligned} e^A &= e^{PDP^{-1}} \\ &= \sum_{k \geq 0} \frac{1}{k!} (PDP^{-1})^k \\ &= \sum_{k \geq 0} \frac{1}{k!} PD^k P^{-1} \\ &= P \left(\sum_{k \geq 0} \frac{1}{k!} D^k \right) P^{-1} \\ &= Pe^D P^{-1}. \end{aligned}$$

- (b)

$$\begin{aligned} \det(e^A) &= \det(Pe^D P^{-1}) \\ &= \det(e^D) \\ &= e^{\lambda_1} e^{\lambda^2} \cdots e^{\lambda_n} \\ &= e^{\lambda_1 + \lambda^2 + \cdots + \lambda_n} \\ &= e^{\text{trace}(A)}. \end{aligned}$$



14. (a) T
 (c) F
 (e) T
 (g) F
 (i) T

Section 19

1. (a) Eigenvalues: 3 and -1 ; $\begin{bmatrix} 1 & -1 \end{bmatrix}^T$ is an eigenvector for A with eigenvalue -1 and $\begin{bmatrix} 1 & 1 \end{bmatrix}^T$ is an eigenvector for A with eigenvalue 3
 (b) As k increases, the vectors $A^k \mathbf{x}_0$ are approaching the vector $\begin{bmatrix} 1 & 1 \end{bmatrix}^T$, which is a dominant eigenvector of A .
 (c) The Rayleigh quotients $r_k = \frac{\mathbf{x}_{k+1}^T \mathbf{x}_k}{\mathbf{x}_k^T \mathbf{x}_k}$ approach the dominant eigenvalue 3.
 (d) Apply the power method to $B = (A - 0I_2)^{-1} = A^{-1}$. As k increases, the vectors $B^k \mathbf{x}_0$ are approaching the vector $\frac{1}{2}\begin{bmatrix} 1 & -1 \end{bmatrix}^T$, which is an eigenvector of A . The Rayleigh quotients approach the other eigenvalue -1 of A .
3. $\begin{bmatrix} 1 & 1 \end{bmatrix}^T$ is an eigenvector for A with eigenvalue 1, so the vectors $A^k \mathbf{x}_0$ are all equal to \mathbf{x}_0 . We can adjust the seed to a non-eigenvector.
5. 8 is the dominant eigenvalue of A
7. (a) Since \mathbb{R}^n has dimension n , it follows that any set of $n + 1$ vectors is linearly dependent.
 (b) Proceed down the list c_{n-1}, c_{n-2} , etc., until you reach a weight that is non-zero.
 (c) $\mathbf{0} = q(A)\mathbf{v} = (A - \lambda I_n)Q(A)\mathbf{v}$
 (d) i. $q(t) = 24t - 10t^2 + t^3$

- ii. 0, 4, and 6
 iii. For $t = 0$, we have $Q(t) = (4 - t)t - 4(t - 6)$, and $[6 \ 6 \ 0]^T$ an eigenvector for A with eigenvalue 0. For $t = 4$ we have $Q(t) = t(t - 6)$, and $[24 \ 24 \ 0]^T$ is an eigenvector for A with eigenvalue 4. For $t = 6$ we have $Q(t) = t(t - 4)$, and $[0 \ -6 \ -12]^T$ is an eigenvector for A with eigenvalue 6.

9. If β is an eigenvalue of B with eigenvector \mathbf{x} , then $\frac{1}{\beta} + \alpha$ is an eigenvalue of A with eigenvector \mathbf{x} .

10. (a) F
 (c) T

Section 20

1. (a) Eigenvalues: $\lambda_1 = 2 + 2\sqrt{2}i$ and $\lambda_2 = 2 - 2\sqrt{2}i$; Eigenvectors: $[-\sqrt{2}i \ 1]^T$ and $[\sqrt{2}i \ 1]^T$
 (b) Eigenvalues: $\lambda_1 = 2+i$ and $\lambda_2 = 2-i$; Eigenvectors: $[-(1+i) \ 1]^T$, $[-(1-i) \ 1]^T$
 (c) Eigenvalues: $\lambda_1 = -1 + 2i$ and $\lambda_2 = -1 - 2i$; Eigenvectors: $[(1+i) \ 2]^T$ and $[(1-i) \ 2]^T$
3. Just the rotation matrices
5. $\begin{bmatrix} 1 & -2 \\ 2 & 1 \end{bmatrix}$
7. (a) Characteristic polynomial $\lambda^2 + a_1\lambda + a_0$; $\begin{bmatrix} 0 & -2 \\ 1 & 2 \end{bmatrix}$
 (b) $\det(C - \lambda I_3) = -\lambda^3 - a_2\lambda^2 - a_1\lambda - a_0$
8. (a) T
 (c) F
 (e) T



Section 21

1. $\det(rA) = r^n \det(A)$

3. (a) $A_1 = \begin{bmatrix} 1 & 1 & 1 & 1 \\ -1 & 4 & -1 & -1 \\ -1 & -1 & 4 & -1 \\ -1 & -1 & -1 & 4 \end{bmatrix}$

(b) $B = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 5 & 0 & 0 \\ 0 & 0 & 5 & 0 \\ 0 & 0 & 0 & 5 \end{bmatrix}$

(c) $125 = \det(B) = \det(A_1) = \det(A)$

(d) $\det(A) = (n+1)^{n-1}$

5. $A^{-1} = \frac{1}{\det(A)} \text{adj}(A)$

$$= \begin{bmatrix} -1 & 0 & 1 \\ 0 & 1 & 0 \\ 2 & 0 & -1 \end{bmatrix}$$

7. (a) 6

(b) -16

9. $|c - 7|$, volume is 0 when the parallelepiped is two-dimensional.

10. (a) T

(c) T

(e) F

(g) T

Section 22

1. (a) Use the fact that $\mathbf{0}_1 + \mathbf{v} = \mathbf{v}$ for any vector \mathbf{v} in our vector space.

(b) Same reasoning as in part (a).

(c) Use the transitive property of equality.

3. Use the fact that $-1 + 1 = 0$.

5. The intersection $W_1 \cap W_2$ is a subspace of V , but the union $W_1 \cup W_2$ is not in general a subspace of V .

7. The space W is the span of $\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$, $\begin{bmatrix} 1 & 0 \\ 3 & 1 \end{bmatrix}$, and $\begin{bmatrix} 0 & -2 \\ 1 & 1 \end{bmatrix}$.

9. Mimic the proof of Theorem 12.5.

11. (a) Closure is by definition, the sequence $\{0\}$ is the additive identity, the sequence $\{-x_n\}$ is the additive inverse of the sequence $\{x_n\}$. The other properties follow from the definitions of addition and multiplication by scalars.

(b) The answer is no.

(c) The answer is yes.

(d) The answer is no.

(e) The answer is yes.

(f) To show that ℓ^2 is closed under addition, expand the square.

13. (a) T

(c) T

(e) T

(g) T

(i) F

Section 23

1. (a) This set is a basis for \mathbb{R}^3 .

(b) The set is linearly independent in \mathbb{P}_3 but does not span \mathbb{P}_3 .

(c) The set is a basis for \mathbb{P}_3 .

(d) The set is linearly independent in $M_{3 \times 2}$ but does not span $M_{3 \times 2}$.

3. The set $\{M_{ij}\}$ where M_{ij} is the matrix with a 1 in the ij th position and zeros everywhere else is a basis for $\mathcal{M}_{2 \times 2}$, as is the set $\{M'_{ij}\}$, where M'_{ij} is the matrix with a -1 in the ij th position and zeros everywhere else.

5. The set is a basis for V .



7. It is not possible.
9. Mimic the proof of Theorem 6.2.
11. Mimic the proof of Theorem 6.4.
- i. Consider the set $\mathcal{S} = \{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3, \mathbf{v}_1\}$.
- ii. Write \mathbf{v}_2 and \mathbf{v}_3 in terms of the vectors in \mathcal{S} .
13. (a) F
(c) T
(e) T
(g) F
(i) F
(k) T
- 11.
- (1) (a) T
(c) T
(e) F
(g) T
(i) F
(k) T

Section 24

1. The set $\{1+t^2, 2+t+2t^2+t^3, 1+t+t^3\}$ is a basis for W and $\dim(W) = 3$.

3. (a) No.

(b) The set S is linearly dependent.

(c) The set $\{A, B, E\}$ forms a basis for $\text{Span } S$ and $C = 2A - 3B$, $D = 3A - 2B$.

(d) Let $F = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}$,
 $G = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$, and
 $H = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$. The set $\{A, B, E, F, G, H\}$ is a basis for $\mathcal{M}_{2 \times 3}$.

5. The set $\mathcal{B} = \left\{ \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \right\}$ is a basis for W . It follows that $\dim(W) = 3$.

7. The set $\mathcal{B} = \{-1+t, -1+t^2\}$ is a basis for W and $\dim(W) = 2$.

9. Consider dimensions.

11. (a) What happens if If \mathbf{v}_i is in W_1 for each i ?

(b) Assume that $\mathbf{v}_1 \notin W_1$.

Section 25

1. $[\mathbf{b}]_{\mathcal{B}} = \begin{bmatrix} -2 \\ 3 \end{bmatrix}$.

3. Two such bases are $\{[1 \ 0]^T, [3 \ 3]^T\}$ and $\{[1 \ 1]^T, [3 \ 1]^T\}$.

5. $\mathcal{B} = \{[1 \ 0 \ 2]^T, [2 \ 1 \ 1]^T\}$.

7. (a) Show that \mathcal{B} is linearly independent.

(b) i. $[p_1(t)]_{\mathcal{B}} = [1 \ 0 \ 1]^T$,
 $[p_2(t)]_{\mathcal{B}} = [1 \ 1 \ 1]^T$, and
 $[p_3(t)]_{\mathcal{B}} = [2 \ -1 \ -1]^T$.

ii. Row reduce the matrix

$$\begin{bmatrix} 1 & 1 & 2 \\ 0 & 1 & -1 \\ 2 & 2 & 1 \end{bmatrix}$$
.

9. (a) Show that \mathcal{B} is linearly independent.

(b) $[A]_{\mathcal{B}} = [0 \ 0 \ 0 \ 1]^T$, $[B]_{\mathcal{B}} = [0 \ 1 \ 1 \ 0]^T$, $[C]_{\mathcal{B}} = [1 \ 1 \ 1 \ 0]^T$, and $[D]_{\mathcal{B}} = [1 \ -1 \ -1 \ 1]^T$.

(c) Row reduce $[[A]_{\mathcal{B}} \ [B]_{\mathcal{B}} \ [C]_{\mathcal{B}} \ [D]_{\mathcal{B}}]$.

11. Use the fact that $c\mathbf{u} = c(u_1\mathbf{v}_1 + u_2\mathbf{v}_2 + \dots + u_n\mathbf{v}_n) = (cu_1)\mathbf{v}_1 + (cu_2)\mathbf{v}_2 + \dots + (cu_n)\mathbf{v}_n$.

13. (a) Write \mathbf{x} as a linear combination of basis vectors.

- (b) Apply T to an appropriate vector and use the fact that T is one-to-one.
15. Find a vector \mathbf{y} in V such that $\mathbf{x} = [\mathbf{y}]_{\mathcal{B}}$.
17. (a) F
(c) F
(e) T
(g) T
(i) T
- (c) $[\mathbf{y}]_{\mathcal{B}} \approx \begin{bmatrix} 0.2 \\ 3.6 \end{bmatrix}$.
7. (a) What is $[\mathbf{v}]_{\mathcal{S}}$ if \mathcal{S} is a standard basis?
(b) Use properties of change of basis matrices.
(c) See the discussion after Activity 26.1.
8. (a) T
(c) T

Section 26

1. (a) $\begin{bmatrix} 5 & -1 & -1 \\ 1 & -1 & 0 \\ -2 & 1 & 1 \end{bmatrix}$.
(b) $\begin{bmatrix} 1 & 1 & 0 & 0 \\ -1 & 1 & 1 & -1 \\ 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix}$.
(c) $\begin{bmatrix} 1 & -1 & 1 & -1 \\ -1 & 1 & 0 & -1 \\ 0 & 0 & -1 & 1 \\ 2 & 0 & 2 & 1 \end{bmatrix}$.

3.

(a) $P_{C \leftarrow \mathcal{B}} = \begin{bmatrix} \frac{5}{2} & -2 \\ \frac{1}{2} & 1 \end{bmatrix}$ and $P_{\mathcal{B} \leftarrow C} = \begin{bmatrix} \frac{2}{7} & \frac{4}{7} \\ -\frac{1}{7} & \frac{5}{7} \end{bmatrix}$.

- (b) i. The columns are linearly independent.
ii. Show that $P_{C \leftarrow \mathcal{B}}^{-1}[\mathbf{x}]_C = [\mathbf{x}]_{\mathcal{B}}$. Then show that $P_{C \leftarrow \mathcal{B}}^{-1}$ and $P_{\mathcal{B} \leftarrow C}$ have the same corresponding columns.

5. (a) Use the trigonometric identities $\cos(\theta + \pi/2) = -\sin(\theta)$ and $\sin(\theta + \pi/2) = \cos(\theta)$.

(b) $[\mathbf{x}]_C \approx \begin{bmatrix} 3.2 \\ 1.6 \end{bmatrix}$.

Section 27

1. (a) The angle between \mathbf{u} and \mathbf{v} is $\frac{\pi}{2}$. The distance between \mathbf{u} and \mathbf{v} is $\sqrt{10}$. The orthogonal projection of \mathbf{u} onto \mathbf{v} is $\mathbf{0}$.
(b) The angle between \mathbf{u} and \mathbf{v} is 0 . The distance between \mathbf{u} and \mathbf{v} is $\sqrt{2}$. $\text{proj}_{\mathbf{v}} \mathbf{u} = \mathbf{u}$.
(c) The angle between \mathbf{u} and \mathbf{v} is approximately 104.96° . The distance between \mathbf{u} and \mathbf{v} is $\sqrt{17}$.
 $\text{proj}_{\mathbf{v}} \mathbf{u} = \frac{\mathbf{u} \cdot \mathbf{v}}{\|\mathbf{v}\|^2} \mathbf{v} = -\frac{1}{10}[1 \ 3]^T$.
(d) The angle between \mathbf{u} and \mathbf{v} is $\frac{\pi}{2}$. The orthogonal projection of \mathbf{u} onto \mathbf{v} is $\mathbf{0}$. The distance between \mathbf{u} and \mathbf{v} is $\sqrt{11}$.
(e) The angle between \mathbf{u} and \mathbf{v} is approximately 54.74° . The distance between \mathbf{u} and \mathbf{v} is $\sqrt{2}$.
 $\text{proj}_{\mathbf{v}} \mathbf{u} = \frac{\mathbf{u} \cdot \mathbf{v}}{\|\mathbf{v}\|^2} \mathbf{v} = \frac{1}{3}[1 \ 1 \ 1]^T$.
3. $h = 6 \pm 4\sqrt{3}$
5. Write the dot product as a matrix-vector product.
7. (a) Write the norm as a dot product.
(b) From part (a), what can we say about $2(\mathbf{u} \cdot \mathbf{v})$?
9. Consider cases of $\mathbf{u} \cdot \mathbf{v} \leq 0$ separately. Then write the norm as a dot product.

11. Use the definition of W^\perp .
13. If $\mathbf{w} \in W_2^\perp$ and $\mathbf{v} \in W_1$, in what other set is \mathbf{v} ?
14. (a) F
 (c) F
 (e) F
 (g) F
 (i) T
 (k) T
 (m) F
 (o) T
 (q) T
- (b) An ellipse centered at the origin with major axis the segment from $(0, -\frac{1}{\sqrt{3}})$ to $(0, \frac{1}{\sqrt{3}})$ and minor axis the segment from $(-\frac{1}{\sqrt{2}}, 0)$ to $(\frac{1}{\sqrt{2}}, 0)$.
5. (a) Verify the inner product properties. Why is the assumption that the a_i are positive necessary?
 (b)
$$\begin{bmatrix} a_1 & 0 & 0 & \cdots & 0 \\ 0 & a_2 & 0 & \cdots & 0 \\ & & \ddots & & \\ 0 & 0 & 0 & \cdots & a_n \end{bmatrix}.$$
7. (a) No.
 (b) If A is a diagonal matrix with positive diagonal entries.

Section 28

1. $\left\{ [0 \ 1 \ 0]^\top, [\frac{3}{4} \ 0 \ 1]^\top \right\}$
3. (a) Where are $\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_k$?
 (b) Take the dot product of \mathbf{z} with \mathbf{w}_i .
 (c) Use parts (a) and (b).
 (d) Collect terms in W and in W^\perp .
5. Consider $(P\mathbf{e}_i) \cdot (P\mathbf{e}_j)$ where \mathbf{e}_t is the t th standard basis vector for \mathbb{R}^n .
7. (a) Simplify $(AB)^\top(AB)$.
 (b) Take the transpose of A^\top .
 (c) What is $(A^{-1})^\top A^{-1}$?
8. (a) F
 (c) T
 (e) F
 (g) T
 (i) T
9. (a) Evaluate each side of the inequality.
 (b) Write $\|\mathbf{w}\|^2$ as an inner product and expand.
11. (a) Expand the inner product.
 (b) Expand the inner product.
 (c) Convert A and B to vectors in \mathbb{R}^{n^2} whose entries are the entries in the first row followed by the entries in the second row and so on.
13. (a) Compute the inner product of the vectors.
 (b) Try to write \mathbf{v} in terms of the basis vectors for W .
 (c) $[1 \ 2 \ 1]^\top$. Approximately 1.41.
15. (a) Use the fact that $\mathbf{0} = \mathbf{0} + \mathbf{0}$.
 (b) Use the fact that $\langle \mathbf{u}, \mathbf{v} \rangle = \langle \mathbf{v}, \mathbf{u} \rangle$.
 (c) Same hint as part (b).
 (d) Use the fact that $\mathbf{u} - \mathbf{v} = \mathbf{u} + (-\mathbf{v})$.

Section 29

1. Use properties of continuous functions.
3. (a) $2x^2 + 3y^2 = 1$
17. Mimic Theorem 28.3.
19. (a) F



- (c) T
(e) F
(g) F
(i) T

19. (a) F
(c) T
(e) T
(g) T

Section 30

1. (a) $\text{proj}_W \mathbf{v} = \frac{1}{2}[1 \ 0 \ 1]^\top$
(b) $\frac{1}{2}[1 \ 0 \ 1]^\top, \sqrt{\frac{1}{2}}$.
3. (a) $\text{proj}_W h(t) \approx -0.386t^2 - 0.721t + 2.06$
(b) $2 - 2t^2$
(c) $\text{proj}_W h(t)$
5. (a) $\left\{ \frac{1}{\sqrt{3}}[1 \ 1 \ 1]^\top, \frac{1}{\sqrt{2}}[1 \ -1 \ 0]^\top \right\}$
(b) $\left\{ \frac{\sqrt{6}}{4}(1+t), \frac{\sqrt{2}}{4}(1 - 3t), \frac{\sqrt{10}}{4}\left(\frac{1}{3} - t^2\right) \right\}$
(c) $\left\{ \frac{1}{\sqrt{14}}[1 \ 0 \ 2]^\top, \frac{\sqrt{35}}{70}[-9 \ 1 \ 3]^\top, \frac{1}{\sqrt{10}}[1 \ 2 \ -2]^\top \right\}$

$$(d) \left\{ \begin{array}{l} \frac{1}{\sqrt{5}} \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \\ 1 \\ 0 \\ 1 \end{bmatrix}, \quad \frac{\sqrt{3}}{6} \begin{bmatrix} -2 \\ 2 \\ 2 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \\ \frac{1}{\sqrt{5}} \begin{bmatrix} 0 \\ -1 \\ 1 \\ -1 \\ 0 \\ 0 \\ -1 \end{bmatrix}, \quad \frac{\sqrt{105}}{210} \begin{bmatrix} 5 \\ 1 \\ 4 \\ 6 \\ 0 \\ 15 \\ -9 \end{bmatrix} \end{array} \right\}$$

7. (a) Evaluate an appropriate equation at well-chosen points.
(b) $\{1, \cos(t), \sin(t) - \frac{2}{\pi}\}$
9. We cannot find a QR factorization for this matrix.

Section 31

1. (a) $\frac{1}{\sqrt{5}} \begin{bmatrix} -2 & 2 \\ 1 & 1 \end{bmatrix}$
(b) $\begin{bmatrix} \frac{1}{\sqrt{3}} & -\frac{2}{\sqrt{6}} & 0 \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} \end{bmatrix}$
(c) A is not orthogonally diagonalizable.
3. $\frac{1}{2} \begin{bmatrix} 9 & 3 & 0 & 0 \\ 3 & 9 & 0 & 0 \\ 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 4 \end{bmatrix}$
5. (a) Determine P_i^\top
(b) Show that every column of P_i is a scalar multiple of \mathbf{u}_i .
(c) Use the orthonormal basis to simplify P_i^2 .
(d) Use the orthonormal basis to simplify $P_i P_j$.
(e) Use the orthonormal basis to simplify $P_i \mathbf{u}_i$.
(f) Use the orthonormal basis to simplify $P_i \mathbf{u}_j$.
(g) Recall that $\text{proj}_{W_i} \mathbf{v} = \frac{\mathbf{v} \cdot \mathbf{u}_i}{\mathbf{u}_i \cdot \mathbf{u}_i} \mathbf{u}_i$.
7. (a) A basis for the eigenspace of A corresponding to the eigenvalue -1 is $\{[0 \ 0 \ -2 \ 1]^\top, [-2 \ 1 \ 0 \ 0]^\top\}$ and a basis for the eigenspace of A corresponding to the eigenvalue 4 is $\{[0 \ 0 \ 1 \ 2]^\top, [1 \ 2 \ 0 \ 0]^\top\}$.
 P_1 is $\frac{1}{5} \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 4 & -2 \\ 0 & 0 & -2 & 1 \end{bmatrix}$,

$$P_2 \text{ is } \frac{1}{5} \begin{bmatrix} 4 & -2 & 0 & 0 \\ -2 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix},$$

$$P_3 \text{ is } \frac{1}{5} \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 2 & 41 \end{bmatrix},$$

$$P_4 \text{ is } \frac{1}{5} \begin{bmatrix} 1 & 2 & 0 & 0 \\ 2 & 4 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix},$$

$$\mu_1 = -1, \mu_2 = 4, Q_1 \text{ is } \frac{1}{5} \begin{bmatrix} 4 & -2 & 0 & 0 \\ -2 & 1 & 0 & 0 \\ 0 & 0 & 4 & -2 \\ 0 & 0 & -2 & 1 \end{bmatrix}, \text{ and}$$

$$Q_2 \text{ is } \frac{1}{5} \begin{bmatrix} 1 & 2 & 0 & 0 \\ 2 & 4 & 0 & 0 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 2 & 4 \end{bmatrix}.$$

- (b) i. Collect matrices with the same eigenvalues.
ii. Use the fact that each P_i is a symmetric matrix.
iii. Use Theorem 31.8.
iv. Use Theorem 31.8.
v. Explain why $\{\mathbf{u}_{1j}, \mathbf{u}_{2j}, \dots, \mathbf{u}_{mj}\}$ is a orthonormal basis for E_{μ_j} .
(c) The rank of Q_j is m_j .

8. (a) T
(c) F
(e) T
(g) T
(i) T
(k) F

Section 32

$$1. \quad (a) \begin{bmatrix} 1 & -1 \\ -1 & 4 \end{bmatrix}$$

$$(b) \begin{bmatrix} 10 & 0 & 2 \\ 0 & 0 & 1 \\ 2 & 1 & 1 \end{bmatrix}$$

$$(c) \begin{bmatrix} 0 & 1 & 1 & -\frac{1}{2} \\ 1 & 5 & 0 & 0 \\ 1 & 0 & 0 & 2 \\ -\frac{1}{2} & 0 & 2 & 8 \end{bmatrix}$$

3. (a) i. Let P be a matrix that orthogonally diagonalizes A , with $P^TAP = D$. Use this to calculate $Q(\mathbf{x})$.
ii. Substitute in part i.
iii. Make an argument similar to part ii.

- (b) The maximum value of $Q(\mathbf{x})$ on the unit circle is 1 and it occurs at the input $\frac{1}{\sqrt{2}}[1 \ -1]$.

5. (a) $1 = \lambda_1 y_1^2 + \lambda_2 y_2^2$.

- (b) An ellipse.
(c) A hyperbola.
(d) Two lines.

7. Use Exercise 6 (a) to compare $Q_A(\mathbf{e}_i)$ and $Q_B(\mathbf{e}_i)$, then compare $Q_A(\mathbf{e}_i + \mathbf{e}_j)$ to $Q_B(\mathbf{e}_i + \mathbf{e}_j)$ for $i \neq j$.

9. (a) F

- (c) T

- (e) T

- (g) F

- (i) T

- (k) T

- (m) T

- (o) T

Section 33

$$1. \quad (a) U = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \Sigma = \begin{bmatrix} \sqrt{2} & 0 \\ 0 & 0 \end{bmatrix}, V = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}.$$



(b) $U = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$, $\Sigma = [\sqrt{2}]$, $V = [1]$. $\Sigma = \begin{bmatrix} \sqrt{28} & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$, $V = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$.

(c) $U = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$, $\Sigma = \begin{bmatrix} \sqrt{3} & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$, $V = \begin{bmatrix} \frac{2}{\sqrt{6}} & 0 & -\frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} \end{bmatrix}$. (c) i. $\left\{ \frac{1}{\sqrt{2}}[-1 \ 1]^T \right\}$.
ii. $\left\{ \frac{1}{\sqrt{14}}[1 \ 2 \ 3]^T \right\}$.
iii. $\left\{ \frac{1}{\sqrt{2}}[1 \ 1]^T \right\}$.

5. Find the transpose of an SVD for A .
7. $\|A\| = \lambda_1$

(d) $U = \begin{bmatrix} \frac{1}{\sqrt{10}} & \frac{1}{\sqrt{10}} & \frac{2}{\sqrt{5}} & 0 \\ \frac{1}{\sqrt{10}} & -\frac{1}{\sqrt{10}} & 0 & -\frac{2}{\sqrt{5}} \\ \frac{2}{\sqrt{10}} & -\frac{2}{\sqrt{10}} & 0 & \frac{1}{\sqrt{5}} \\ \frac{2}{\sqrt{10}} & \frac{2}{\sqrt{10}} & -\frac{2}{\sqrt{5}} & 0 \end{bmatrix}$, 9. Mimic Exercise 5 in Section 31.
9. (a) F
(c) T
(e) F
(g) F

$\Sigma = \begin{bmatrix} 5 & 0 \\ 0 & \sqrt{5} \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$, $V = \text{Section 34}$

1. (a) 60, 15, and 6
(b) $A = 60 \left(\frac{1}{5}[3 \ 0 \ 4]^T \right) \left(\frac{1}{3}[2 \ 1 \ 1] \right) + 15 \left(\frac{1}{5}[4 \ 0 \ -3]^T \right) \left(\frac{1}{3}[-1 \ -2 \ 2] \right) + 6 \left(\frac{1}{5}[0 \ 5 \ 0]^T \right) \left(\frac{1}{3}[-2 \ 2 \ 1] \right)$.

(e) $U = \begin{bmatrix} 0 & 1 & 0 \\ \frac{1}{\sqrt{2}} & 0 & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \end{bmatrix}$, (c) $\begin{bmatrix} 24 & 12 & 24 \\ 0 & 0 & 0 \\ 32 & 16 & 32 \end{bmatrix}$, $\frac{261}{3861} \approx 0.068$.

$\Sigma = \begin{bmatrix} 3 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$, $V = \begin{bmatrix} 20 & 4 & 32 \\ 0 & 0 & 0 \\ 35 & 22 & 26 \end{bmatrix}$, $\frac{36}{3861} \approx 0.009$.

3. (a) $f(x) \approx 35.5273 - 0.0527x$.
(b) Approximately 32.10.
5. (a) The plot demonstrates that the data is not linear.
(b) $g(x) = \frac{266}{97} + \frac{12213}{970}x - \frac{1423}{9700}x^2$.

3. (a) $\sqrt{28}$

(b) $U = \begin{bmatrix} \frac{1}{\sqrt{14}} & \frac{2}{\sqrt{5}} & \frac{3}{\sqrt{70}} \\ \frac{2}{\sqrt{14}} & -\frac{1}{\sqrt{5}} & \frac{6}{\sqrt{70}} \\ \frac{3}{\sqrt{14}} & 0 & -\frac{5}{\sqrt{70}} \end{bmatrix}$, 7. (a) $\det(A) = 1$, eigenvalue 1
(b) i. (2, 1, 1)
ii. (4, 2, 1, 1)



iii. $\mathbf{x} = [x_1 \ x_2 \ \dots \ x_n]^T$ with
 $x_1 = 2^{n-2}, x_2 = 2^{n-3}, \dots,$
 $x_{n-2} = 2, x_{n-1} = 1$, and
 $x_n = 1$

iv. $\|A - B\|_F = \frac{1}{2^{n-2}}$

9. (a) Use the fact that $\Sigma\Sigma^+$ is a symmetric matrix.

- (b) Take the transpose of A^+A .

11. $A^+ = 0$

13. If $A^T A \mathbf{x} = \mathbf{0}$, what is $\mathbf{x}^T A^T A \mathbf{x}$?

14. (a) F

- (c) F

- (e) F

Section 35

1. Use properties of the definite integral from calculus.

3. The property that $T^{-1}(\mathbf{w}) = \mathbf{v}$ whenever $T(\mathbf{v}) = \mathbf{w}$ is the key to this problem.

5. (a) linear transformation, one-to-one, onto

- (b) linear transformation, one-to-one, onto

7. Use the fact that T is one-to-one to show that is linearly independent, and that T is onto to show that \mathcal{C} spans W .

9. Yes, but the vector space needs to be infinite dimensional.

11. Show that $T(\mathbf{v}) = \mathbf{w}$ has at most one solution for each \mathbf{w} in W .

13. (a) Think of $T(V')$ as the range of a suitable restriction of T .

- (b) Use Exercise 12. It is possible.

- (c) Let $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k\}$ be a basis for $\text{Ker}(T)$. Extend this basis to a basis $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k, \mathbf{v}_{k+1}, \dots, \mathbf{v}_n\}$ of V . Use this basis to find a basis for $\text{Range}(T)$.

- (d) This follows from fact that coordinate transformations are linear.

18. (a) F

- (c) T

- (e) F

- (g) T

- (i) T

Section 36

1. (a) $r(t) = r_0(1) + r_1(t) + r_2(t^2)$ and

$$[r(t)]_{\mathcal{B}} = \begin{bmatrix} r_0 \\ r_1 \\ r_2 \end{bmatrix}.$$

- (b) T is a linear transformation

- (c) The coordinate mapping is a linear transformation.

$$(d) [T(p_0(t))]_{\mathcal{C}} = \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix},$$

$$[T(p_1(t))]_{\mathcal{C}} = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix},$$

$$[T(p_2(t))]_{\mathcal{C}} = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 1 \end{bmatrix}$$

$$(e) \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

$$(f) [T(1 + t - t^2)]_{\mathcal{C}} = \begin{bmatrix} 2 \\ 0 \\ 1 \\ -1 \end{bmatrix},$$

$$T(1 + t - t^2) = 2(1) + 0(t) + 1(t^2) + (-1)t^3 = 2 + t^2 - t^3.$$

$$(g) T(1 + t - t^2) = 2 + t^2 - t^3.$$

3. (a) Use the linearity of both S and T .



- (b) True
5. $\mathbf{x} \in \text{Ker}(T)$ if and only if $[\mathbf{x}]_{\mathcal{B}} \in \text{Nul } [T]_{\mathcal{B}}^{\mathcal{C}}$
T
7. (a) Since \mathbf{w} is in the range of T , there is a vector \mathbf{v} so that $T(\mathbf{v}) = \mathbf{w}$.
(b) Since \mathbf{y} is in the range of T' , there exists a vector \mathbf{x} in \mathbb{R}^n so that $T'(\mathbf{x}) = \mathbf{y}$. What is $[\mathbf{w}]_{\mathcal{C}}$?
(c) How do we tell if T' is onto?
9. (a) F
(c) T
(e) T
(g) T
(i) T
- (c) $-1, -2$, and -3 with bases $\{1\}$, $\{-1 + t\}$, $\{1 - 2t + t^2\}$
(d) $\mathcal{B} = \{-1, -1 + t, 1 - 2t + t^2\}$
3. (a) Use properties of differentiable functions.
(b) Use properties of the derivative.
(c) What is $D(e^{\lambda x})$?
5. (a) Use properties of the matrix transpose.
(b) For which matrices is $T(A) = A$?
(c) When is it possible to have $A^T = \lambda A$?

Section 37

1. (a) Use properties of the derivative.

$$(b) [T]_{\mathcal{S}} = \begin{bmatrix} -1 & 1 & 0 \\ 0 & -2 & 2 \\ 0 & 0 & -3 \end{bmatrix}.$$

6. (a) F

- (c) T
(e) T
(g) T

Index

- LU* factorization, 377
angle between vectors, 482, 513
attractor, 649
augmented matrix, 22

back substitution, 26
back-face culling, 491
basic variable, 28
basis, 115
basis for a vector space, 412
Binet's formula, 329

Cauchy-Schwarz inequality
 in \mathbb{R}^n , 490
change of basis matrix, 461
characteristic equation, 299
characteristic polynomial, 299
characteristic value, 170
characteristic vector, 170
circuits
 junctions, 4
 resistor, 3
 source, 3
codomain of a matrix transformation, 128
coefficient matrix, 22
cofactor, 283
column space, 235
companion matrix, 356
complex conjugate, 685
complex number, 683
 imaginary part, 683
 real part, 683
condition number of a matrix, 611
contraction, 140
coordinate transformation, 447

coordinate vector with respect to a basis, 446
coordinates with respect to a basis, 446
covariance matrix, 438
Cramer's Rule, 371
cross product, 492

determinant, 283
diagonal matrix, 153
diagonal of a matrix, 153
differential equation, 672
dimension
 subspace of \mathbb{R}^n , 265
distance between vectors, 481
distance between vectors in an inner product
 space, 512
domain of a matrix transformation, 128
dominant eigenvalue, 333
dominant eigenvector, 333
dot product, 150, 479
dynamical system, 173

eigenspace, 251
eigenvalue, 170
eigenvector, 170
elementary operations, 7
equivalent statements, 203

Fibonacci sequence, 314
field, 59
forward elimination, 26
Fourier coefficients, 529
free variable, 28

Gaussian quadrature, 542
Gershgorin Disk Theorem, 360
Gram-Schmidt Process
 in an inner product space, 533

- Hamming code, 405
 Hausdorff dimension, 652
 homogeneous system, 82
 identity matrix, 152
 image of an element under a transformation, 128
 inner product, 510
 inner product space, 510
 input-output models, 97
 Invertible Matrix Theorem, 204
 isometry, 500
 isomorphism, 642
 iterated function system, 649
 kernel, 233
 Kirchoff's Current Law, 17
 Kirchoff's Voltage Law, 17
 Lagrange polynomials, 667
 Latent Semantic Indexing, 581
 leading entry of a row, 42
 length of a vector in an inner product space, 512
 Leslie matrix, 344
 linear combination, 65
 weights, 65
 linear combination of vectors in a vector space, 396
 linear dependence, 107
 linear dependent vectors in a vector space, 409
 linear equation, 6
 coefficients, 6
 linear independence, 107
 linear transformation, 638
 diagonalizable, 673
 eigenvalue, 672
 eigenvector, 672
 kernel, 640
 matrix of, 658
 onto, 639
 range, 641
 standard matrix, 657
 linearly independent vectors in a vector space, 409
 lower triangular matrix, 154
 Markov chain, 180, 181
 Markov process, 179, 180
 matrix, 22
 adjugate, 370
 column, 23
 columns, 83
 diagonalizable, 319
 elementary, 365
 entry, 23
 exponential, 162
 indefinite, 570
 inverse, 192
 invertible, 192
 minor, 282
 negative definite, 570
 negative semidefinite, 570
 nilpotent, 256
 non-singular, 192
 orthogonal, 500
 positive definite, 570
 positive semidefinite, 570
 product, 147
 row, 23
 rows, 83
 scalar multiple, 144
 singular, 192
 size, 23, 83
 sparse, 335
 square, 171
 sum, 146
 transpose, 153
 matrix transformation, 126
 matrix-vector product, 84
 minimal spanning set, 107
 moment of inertia, 577
 multiplicity
 algebraic, 300
 geometric, 302
 nonhomogeneous system, 82
 normal vector, 484
 null space, 233
 nullity of a matrix, 266
 Ohm's Law, 17
 one-to-one, 131
 onto, 130



- operator norm of a matrix, 582
orthogonal, 482
orthogonal basis in \mathbb{R}^n , 496
orthogonal basis in an inner product space,
 514
orthogonal complement, 484
orthogonal complement in inner product
 spaces, 519
orthogonal diagonalization, 550
orthogonal projection
 in the direction of a vector, 484
orthogonal projection onto a subspace, 516
orthogonal set in \mathbb{R}^n , 496
orthogonal set in an inner product space, 513
orthogonal vectors in inner product spaces,
 513
orthonormal basis in \mathbb{R}^n , 499
orthonormal basis in an inner product space,
 515
outer product decomposition, 590
parametric vector form, 82
partitioned matrices, 164
pivot, 28, 42
pivot column, 42
pivot positions, 42
polynomial curve fitting, 35
power method, 333
projection
 orthogonal to a vector, 484
projection orthogonal to a subspace, 516
pseudoinverse, 614
Pythagorean Theorem in \mathbb{R}^n , 489
quadratic form, 567
range of a matrix transformation, 128
rank of a matrix, 266
Rayleigh quotients, 335
rotation matrix, 138
rotation-scaling matrices, 352
row echelon form, 41
 reduced, 43
row equivalent matrices, 46
row operations, 24
row space, 235
scalar product, 479
scalars, 59, 391
self-similar set, 650
shear, 139
similar matrices, 316
singular values, 587
singular vectors
 left, 589
 right, 588
span, 67
span of a set of vectors in a vector space, 396
spanning set, 396
spectral decomposition, 556
spectrum of a matrix, 554
standard basis
 for \mathbb{P}_n , 412
 for \mathbb{R}^n , 115
state vector, 174
stochastic matrix, 184
 regular, 186
Strassen's algorithm, 163
strictly diagonally dominant matrix, 360
subspace
 of \mathbb{R}^n , 217
 sum, 222, 402
subspace of a vector space, 395
subspace of a vector space spanned by a set
 of vectors, 396
symmetric matrix, 154
system of linear equations, 6
 consistent, 8
 elimination, 7
 equivalent systems, 7
 inconsistent, 8
 operations, 7
 solution, 7
 solution set, 7
trace, 325
transformation
 matrix, 126
transition matrix, 177, 180
triangle inequality in \mathbb{R}^n , 490
trigonometric polynomial, 527
unit vector, 480
unit vector in an inner product space, 512
upper triangular matrix, 154



- vector, 58
 - column, 59
 - component, 58
 - entry, 58
 - length in \mathbb{R}^n , 63, 479
 - magnitude in \mathbb{R}^n , 479
 - norm in \mathbb{R}^n , 479
- vector space, 216, 391
- basis, 412
- dimension, 428
- finite dimensional, 428
- subspace, 395
- vector spaces
 - isomorphic, 642
- Wronskian, 681

